(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 August 2005 (25.08.2005)

PCT

(10) International Publication Number
WO 2005/077403 A1

(51) International Patent Classification[7]: A61K 38/18

(21) International Application Number:
PCT/US2005/003229

(22) International Filing Date: 4 February 2005 (04.02.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/542,260 4 February 2004 (04.02.2004) US

(71) Applicant *(for all designated States except US)*: **STRYKER CORPORATION** [US/US]; 2725 Fairfield Road, Kalamazoo, Michigan 49005 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **LEE, John C.** [US/US]; 1119 Haltown Drive, San Antonio, Texas 78213 (US). **YEH, Lee-Chuan C.** [US/US]; 8027 Indian Bend, San Antonio, Texas 78250 (US).

(74) Agents: **HALEY, James F. Jr.** et al.; FISH & NEAVE IP GROUP, ROPES & GRAY LLP, 1251 Avenue of the Americas, New York, New York 10020 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

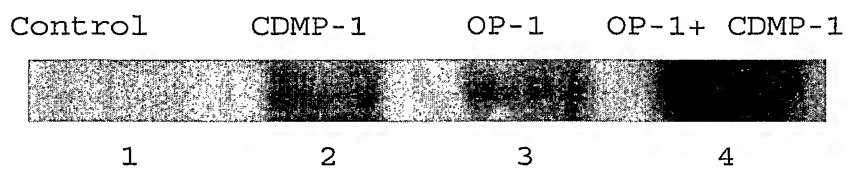
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: THERAPEUTIC METHODS USING SMADS



(57) Abstract: Methods of inducing the expression of a Smad in a cell or tissue comprising the step of contacting the cell or tissue capable of expressing the Smad with a bone morphogenic protein are provided. Methods of inducing tissue formation and repairing a tissue defect or regenerating tissue, at a target locus in a mammal comprising the step of administering to the target locus a Smad are also provided.

WO 2005/077403 A1

Therapeutic Methods Using Smads

Technical Field of the Invention

[0001] The present invention relates to methods for tissue formation, repair and regeneration using Smads.

Background of the Invention

[0002] The TGF-ß superfamily represents a large number of evolutionarily conserved proteins with diverse activities involved in growth, differentiation, cell migration, development, apoptosis and tissue morphogenesis and repair. This large family includes the bone morphogenic proteins (BMPs), TGF-ßs and activins. Each subgroup of proteins initiates a unique signaling cascade activated by the formation of a complex with a receptor. The receptors for this family of proteins are type I also known as activin receptor-like kinases (ALKs) and type II serine/threonine kinases. Several such receptors have been identified thus far. The type II receptors are constitutively active. Upon ligand binding the type II receptor phosphorylates particular serine and threonine residues in the type I receptor. The type I serine/threonine

kinases become activated and transduce signals to downstream molecules.

[0003] The downstream molecules in the signaling cascade include the Smads. These molecules are vertebrate counterparts to the *Drosophila* and *Caenorhabditis elegans*, proteins known as Mad (*mothers against dpp)* and Sma, respectively. The name Smad originates from a fusion between Mad and Sma. In recent years, several Smads have been identified (Smad 1-8) (Derynck R. et al., 1996, "Nomenclature: Vertebrate mediators of TGF-ß family signals", Cell, 18, 173). The Smads can be divided into three groups: receptor-regulated Smads (R-Smads), common-partner Smads (Co-Smads) and inhibitory Smads (I-Smads). R-Smads are bound to the cell membrane through membrane bound proteins. They transiently interact with and are activated by phosphorylation by activated type I receptor kinases. R-Smads include Smad1, Smad2, Smad3, Smad5 and Smad8. Of those, Smad2 and Smad3 are TGF-β- and activin-specific, whereas Smad1, Smad5 and Smad8 are BMP-specific. The activated R-Smads recruit Co-Smads and these heteromeric complexes translocate to the nucleus. Co-Smads include Smad4. These nuclear Smad complexes bind to DNA directly or indirectly through other DNA-binding proteins, and regulate the transcription of target genes. I-Smads interact with the activated type I receptor, and prevent R-Smads from interacting with activated type I receptors. I-Smads include Smad6 and Smad7.

[0004] As described above, the TGF-ß superfamily of proteins have important roles in various physiological events including the inductive properties of the proteins belonging to the BMP family. Therefore, there

remains a need for identifying means useful for promoting tissue regeneration in patients with traumas caused, for example, by injuries or degenerative disorders.

[0005] The ability to induce Smad protein expression in sufficient quantities at a target locus would be very useful in orthopedic medicine, certain types of plastic surgery, dental and various periodontal and craniofacial reconstructive procedures, and procedures generally involving bone, cartilage, tendon, ligament and neural regeneration. Several Smad genes are now cloned and may be recombinantly expressed in a variety of host systems. The ability to recombinantly produce active Smads, including variants and fragments thereof, and to express them at a target locus makes potential therapeutic treatments using these proteins either alone or together with BMPs feasible.

Summary of the Invention

[0006] This invention is based on the discovery that Smad expression is induced in the presence of various bone morphogenic proteins such as OP-1 (BMP-7) and CDMPs. Therefore, this invention provides a method of inducing the expression of a Smad in a cell or tissue comprising the step of contacting the cell or tissue capable of expressing a Smad with a bone morphogenic protein (BMP). In some embodiments, the tissue is selected from the group consisting of bone, cartilage, tendon, ligament and neural tissue. In one preferred embodiment, the tissue is bone or cartilage. In another preferred embodiment, the tissue is tendon or ligament. In a more preferred embodiment, the tissue is bone. In

another more preferred embodiment, the tissue is cartilage.

[0007] In some embodiments, the cells used in the methods of this invention are progenitor cells. In some embodiments, the progenitor cells include an osteoprogenitor cell, a cartilage progenitor cell, a ligament progenitor cell, a tendon progenitor cell, or a neural progenitor cell. In a preferred embodiment, the progenitor cell is an osteoprogenitor cell or a cartilage progenitor cell. In another preferred embodiment, the progenitor cell is a tendon progenitor cell or a ligament progenitor cell. In a more preferred embodiment, the progenitor cell is an osteoprogenitor cell. In another more preferred embodiment, the progenitor cell is a cartilage progenitor cell.

[0008] In some embodiments, the cell or tissue is contacted with more than one BMP. In some embodiments, the cell or tissue is contacted with two BMPs. BMPs include, but are not limited to, OP-1 (BMP-7), OP-2, OP-3, COP-1, COP-3, COP-4, COP-5, COP-7, COP-16, BMP-2, BMP-3, BMP-3b, BMP-4, BMP-5, BMP-6, BMP-9, BMP-10, BMP-11, CDMP-3 (BMP-12), CDMP-2 (BMP-13), CDMP-1 (BMP-14), BMP-15, BMP-16, BMP-17, BMP-18, GDF-1, GDF-2, GDF-3, GDF-5, GDF-6, GDF-7, GDF-8, GDF-9, GDF-10, GDF-11, GDF-12, MP121, dorsalin-1, DPP, Vg-1, Vgr-1, 60A protein, NODAL, UNIVIN, SCREW, ADMP, and NEURAL. In one preferred embodiment, the BMP is OP-1 (BMP-7). In another preferred embodiment, the BMP is CDMP-1 or GDF-5. In yet another preferred embodiment, the first bone morphogenic protein is OP-1 and the second bone morphogenic protein is CDMP-1 or GDF-5.

[0009] In some embodiments, the Smads used in the methods of the present invention include Smad1, Smad2,

Smad3, Smad5 and Smad8. In a preferred embodiment, the Smad is Smad5. In another preferred embodiment, the Smad is a recombinant Smad.

[0010] In some embodiments, the cell or tissue used in the method of inducing the expression of a Smad in a cell or tissue is further capable of expressing a serine/threonine kinase receptor. In some embodiments the serine/threonine kinase receptor is selected from the group consisting of type I and type II receptors. In some embodiments, only type I receptors are used. In some embodiments, only type II receptors are used. In some embodiments, both type I and type II receptors are used. Preferably, the type I and type II receptors are recombinant. In some embodiments the serine/threonine kinase receptors are linked to an expression control sequence. In some embodiments, the expression control sequence comprises a constitutive promoter. In some embodiments, the expression control sequence comprises an inducible promoter. In some embodiments, the type I receptor is an activin receptor-like kinase (ALK). ALKs include but are not limited to ALK-1, ALK-2, ALK-3, ALK-4, ALK-5, ALK-6, ALK-7 and fragments thereof.

[0011] The invention also provides gene therapy methods of inducing tissue formation, repairing a tissue defect or regenerating tissue at a target locus. In some embodiments, the invention provides a method of inducing tissue formation at a target locus in a mammal comprising the step of administering to the target locus a nucleic acid encoding a Smad. In some embodiments, the invention provides a method of inducing tissue formation at a target locus in a mammal comprising the step of administering to the target locus a vector comprising a nucleic acid encoding a Smad operably

linked to an expression control sequence. In some embodiments, the invention provides a method of inducing tissue formation at a target locus in a mammal comprising the step of administering to the target locus a cell comprising a vector comprising a nucleic acid encoding a Smad operably linked to an expression control sequence.

[0012] The invention further provides a method of repairing a tissue defect or regenerating tissue at a target locus in a mammal comprising the step of administering to the target locus a nucleic acid encoding a Smad. In some embodiments, the invention provides a method of repairing a tissue defect or regenerating tissue at a target locus in a mammal comprising the step of administering to the target locus a vector comprising a nucleic acid encoding a Smad operably linked to an expression control sequence. In some embodiments, the invention provides a method of repairing a tissue defect or regenerating tissue at a target locus in a mammal comprising the step of administering to the target locus a cell comprising a vector comprising a nucleic acid encoding a Smad operably linked to an expression control sequence.

[0013] In some embodiments, the tissue in the gene therapy methods of the present invention is selected from the group consisting of bone, cartilage, tendon, ligament and neural tissue. In one preferred embodiment, the tissue is bone or cartilage. In another preferred embodiment, the tissue is tendon or ligament. In a more preferred embodiment, the tissue is bone. In another more preferred embodiment, the tissue is cartilage.

[0014] In some embodiments, the cells used in the gene therapy methods of this invention are progenitor cells. In some embodiments, the progenitor cells include an osteoprogenitor cell, a cartilage progenitor cell, a ligament progenitor cell, a tendon progenitor cell, or a neural progenitor cell. In a preferred embodiment, the progenitor cell is an osteoprogenitor cell or a cartilage progenitor cell. In another preferred embodiment, the progenitor cell is a tendon progenitor cell or a ligament progenitor cell. In a more preferred embodiment, the progenitor cell is an osteoprogenitor cell. In another more preferred embodiment, the progenitor cell is a cartilage progenitor cell.

[0015] In some embodiments, the expression control sequence operably linked to the nucleic acid encoding a Smad comprises a constitutive promoter. In other embodiments, the expression control sequence operably linked to the nucleic acid encoding a Smad comprises an inducible promoter. A Smad according to this invention includes, but is not limited to, Smad1, Smad2, Smad3, Smad5, Smad8 and fragments thereof. In a preferred embodiment, the Smad is Smad5. In another preferred embodiment, the Smad is a recombinant Smad.

[0016] In some embodiments, the methods of inducing tissue formation, repairing a tissue defect or regeneration tissue of this invention further comprise the step of administering to the target locus a serine/threonine kinase receptor. In some embodiments, a nucleic acid encoding a serine/threonine kinase receptor is administered. In some embodiments, a vector comprising a nucleic acid encoding a serine/threonine kinase receptor operably linked to an expression control

sequence is administered. In other embodiments, a cell comprising a vector comprising a nucleic acid encoding a serine/threonine kinase receptor operably linked to an expression control sequence is administered.

[0017] In some embodiments, the expression control sequence operably linked to the a serine/threonine kinase receptor comprises a constitutive promoter. In some embodiments, the expression control sequence operably linked to the serine/threonine kinase receptor comprises an inducible promoter.

[0018] In some embodiments, the serine/threonine kinase receptor used in the gene therapy methods of inducing tissue formation, repairing a tissue defect or regenerating tissue at a target locus, is selected from the group consisting of type I and type II receptors. In some embodiments, only type I receptors are used. In some embodiments, only type II receptors are used. In some embodiments both type I and type II receptors are used. Preferably recombinant type I and recombinant type II receptors are used. In some embodiments, the type I receptor is an activin receptor-like kinase (ALK). The ALKs that may be used in the present invention include, but are not limited to ALK-1, ALK-2, ALK-3, ALK-4, ALK-5, ALK-6, ALK-7, and fragments thereof. In a preferred embodiment, the ALK is ALK-2, ALK-3, ALK-6 and fragments thereof.

[0019] In some embodiments, the methods of inducing tissue formation, repairing a tissue defect or regeneration tissue of this invention further comprise the step of administering to the target locus a bone morphogenic protein. In some embodiments, the bone morphogenic protein is administered as a nucleic acid. In some embodiments, the bone morphogenic protein is

administered as a vector comprising a nucleic acid encoding the bone morphogenic protein operably linked to an expression control sequence. In some embodiments, the bone morphogenic protein is administered as a cell comprising a vector comprising a nucleic acid encoding a bone morphogenic protein operably linked to an expression control sequence. In some embodiments, the expression control sequence linked to the BMP nucleic acid comprises a constitutive promoter. In some embodiments, the expression control sequence linked to the BMP nucleic acid comprises an inducible promoter.

[0020] The bone morphogenic protein according to this invention includes, but is not limited to, OP-1 (BMP-7), OP-2, OP-3, COP-1, COP-3, COP-4, COP-5, COP-7, COP-16, BMP-2, BMP-3, BMP-3b, BMP-4, BMP-5, BMP-6, BMP-9, BMP-10, BMP-11, CDMP-3 (BMP-12), CDMP-2 (BMP-13), CDMP-1 (BMP-14), BMP-15, BMP-16, BMP-17, BMP-18, GDF-1, GDF-2, GDF-3, GDF-5, GDF-6, GDF-7, GDF-8, GDF-9, GDF-10, GDF-11, GDF-12, MP121, dorsalin-1, DPP, Vg-1, Vgr-1, 60A protein, NODAL, UNIVIN, SCREW, ADMP, NEURAL, and fragments thereof. In some embodiments, one BMP is administered. In some embodiments, more than one BMP is administered. In some embodiments, two BMPs are administered. In some embodiments, the first BMP is OP-1 and the second BMP is CDMP-1 or GDF-5. In a preferred embodiment, the BMP is OP-1. In another preferred embodiment, the BMP is CDMP-1 or GDF-5.

Brief Description of the Drawings

[0021] Figure 1 is a Western blot analysis for Smad5 protein in C2C12 cells following treatment with OP-1 and CDMP-1. Lane 1 is the Smad5 levels in control C2C12 cells. Lane 2 is the Smad5 levels in CDMP-1 treated

(200 ng/ml) C2C12 cells. Lane 3 is the Smad5 levels in OP-1 treated (100 ng/ml) C2C12 cells. Lane 4 is the Smad5 levels in cells treated with both OP-1 (100 ng/ml) and CDMP-1 (200 ng/ml).

[0022] Figure 2 is a plasmid map of pW24 containing the OP-1 gene.

Detailed Description of the Invention

[0023] In order that the invention herein described may be fully understood, the following detailed description is set forth.

[0024] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as those commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. The materials, methods and examples are illustrative only, and are not intended to be limiting. All publications, patents and other documents mentioned herein are incorporated by reference in their entirety.

[0025] Throughout this specification, the word "comprise" or variations such as "comprises" or "comprising" will be understood to imply the inclusion of a stated integer or groups of integers but not the exclusion of any other integer or group of integers.

[0026] In order to further define the invention, the following terms and definitions are provided herein.

[0027] The term "defect" or "defect site," refers to a disruption of the specified tissue. A defect can assume the configuration of a "void", which is

understood to mean a three-dimensional defect such as, for example, a gap, cavity, hole or other substantial disruption in the structural integrity of the tissue (e.g., bone, chondral, osteochondral, neural, ligament, tendon). Moreover, a defect can also be a detachment of the tendon or ligament from its point of attachment to bone, cartilage or muscle. In certain embodiments, the defect is such that it is incapable of endogenous or spontaneous repair. A defect can be the result of accident, disease, and/or surgical manipulation.

[0028] The term "target locus" refers to the site (e.g., a defect) in any tissue that is in need of repair or regeneration. The target locus need not be a defect sit. It may be any site where bone, cartilage, tendon, ligament or neural tissue regeneration is desired.

[0029] The term "repair" refers to new tissue formation which is sufficient to at least partially fill the void or structural discontinuity at the defect site. Repair does not, however, mean, or otherwise necessitate, a process of complete healing or a treatment, which is 100% effective at restoring a defect to its pre-defect physiological/structural/ mechanical state.

[0030] The term "therapeutically effective amount" refers to an amount effective to repair, regenerate, promote, accelerate, prevent degradation, or form tissue.

[0031] The term "patient" refers to an animal, including a mammal (e.g., a human).

[0032] The term "bone morphogenic protein (BMP)" refers to a protein belonging to the BMP family of the TGF-β superfamily of proteins (BMP family) based on DNA and amino acid sequence homology. A protein belongs to

the BMP family according to this invention when it has at least 50% amino acid sequence identity with at least one known BMP family member within the conserved C-terminal cysteine-rich domain, which characterizes the BMP protein family. Preferably, the protein has at least 70% amino acid sequence identity with at least one known BMP family member within the conserved C-terminal cysteine rich domain. Members of the BMP family may have less than 50% DNA or amino acid sequence identity overall. Bone morphogenic proteins may be monomeric, homo- or hetero-dimeric. Bone morphogenic proteins include osteogenic proteins.

[0033] Bone morphogenic proteins are capable of inducing progenitor cells to proliferate and/or to initiate differentiation pathways that lead to cartilage, bone, tendon, ligament or other types of tissue formation depending on local environmental cues, and thus bone morphogenic proteins may behave differently in different surroundings. For example, a bone morphogenic protein may induce bone tissue at one treatment site and cartilage tissue at a different treatment site. Bone morphogenic proteins include full length proteins as well as fragments thereof.

[0034] The term "osteogenic protein (OP)" refers to a bone morphogenic protein that is capable of inducing a progenitor cell to form cartilage and/or bone. The bone may be intramembranous bone or endochondral bone. Most osteogenic proteins are members of the BMP protein family and are thus also BMPs. As described elsewhere herein, the class of proteins is typified by human osteogenic protein (hOP-1). Other osteogenic proteins useful in the practice of the invention include but are not limited to, osteogenically active forms of OP-1, OP-

2, OP-3, COP-1, COP-3, COP-4, COP-5, COP-7, COP-16, BMP-2, BMP-3, BMP-3b, BMP-4, BMP-5, BMP-6, BMP-9, BMP-10, BMP-11, CDMP-3 (BMP-12), CDMP-2 (BMP-13), CDMP-1 (BMP-14), BMP-15, BMP-16, BMP-17, BMP-18, GDF-1, GDF-2, GDF-3, GDF-5, GDF-6, GDF-7, GDF-8, GDF-9, GDF-10, GDF-11, GDF-12, MP121, dorsalin-1, DPP, Vg-1, Vgr-1, 60A protein, NODAL, UNIVIN, SCREW, ADMP, NEURAL, conservative amino acid sequence variants thereof having osteogenic activity and fragments thereof. In one currently preferred embodiment, osteogenic protein includes any one of: OP-1, CDMP-1, CDMP-2, CDMP-3, GDF-5, GDF-6, GDF-7, amino acid sequence variants and homologs thereof, including species homologs thereof and fragments thereof. Particularly preferred osteogenic proteins are those comprising an amino acid sequence having at least 70% homology with the C-terminal 102-106 amino acids, defining the conserved seven cysteine domain, of, e.g., human OP-1. Certain preferred embodiments of the instant invention comprise the osteogenic protein, OP-1. As further described elsewhere herein, the osteogenic proteins suitable for use with this invention can be identified by means of routine experimentation using the art-recognized bioassay described by Reddi and Sampath (Sampath *et al.*, *Proc. Natl. Acad. Sci.*, 84, pp. 7109-13, incorporated herein by reference).

[0035] Proteins useful in this invention include eukaryotic proteins identified as osteogenic proteins (see U.S. Patent 5,011,691, incorporated herein by reference), such as the OP-1, OP-2, OP-3 and CBMP-2 proteins, as well as amino acid sequence-related proteins, such as DPP (from Drosophila), Vg1 (from Xenopus), Vgr-1 (from mouse), GDF-1 (from humans, see

Lee, *PNAS*, 88, pp. 4250-4254 (1991)), 60A (from Drosophila, see Wharton *et al.*, *PNAS*, 88, pp. 9214-9218 (1991)), dorsalin-1 (from chick, see Basler *et al.*, *Cell*, 73, pp. 687-702 (1993) and GenBank accession number L12032) and GDF-5 (from mouse, see Storm *et al.*, *Nature*, 368, pp. 639-643 (1994)). The teachings of the above references are incorporated herein by reference. BMP-3 is also preferred. Additional useful proteins include biosynthetic morphogenic constructs disclosed in U.S. Pat. No. 5,011,691, incorporated herein by reference, e.g., COP-1, COP-3, COP-4, COP-5, COP-7 and COP-16, as well as other proteins known in the art. Still other proteins include osteogenically active forms of BMP-3b (see Takao, *et al.*, *Biochem. Biophys. Res. Comm.*, 219, pp. 656-662 (1996)). BMP-9 (see WO 95/33830), BMP-15 (see WO 96/35710), BMP-12 (see WO 95/16035), CDMP-1 (see WO 94/12814), CDMP-2 (see WO 94/12814), BMP-10 (see WO 94/26893), GDF-1 (see WO 92/00382), GDF-10 (see WO95/10539), GDF-3 (see WO 94/15965) and GDF-7 (see WO95/01802). The teachings of the above references are incorporated herein by reference. BMPs (identified by sequence homology) must have demonstrable osteogenic activity in a functional bioassay to be osteogenic proteins according to this invention.

[0036] The term "amino acid sequence homology" is understood to include both amino acid sequence identity and similarity. Homologous sequences share identical and/or similar amino acid residues, where similar residues are conservative substitutions for, or "allowed point mutations" of, corresponding amino acid residues in an aligned reference sequence. Thus, a candidate polypeptide sequence that shares 70% amino acid homology

with a reference sequence is one in which any 70% of the aligned residues are either identical to, or are conservative substitutions of, the corresponding residues in a reference sequence. Certain particularly preferred bone morphogenic polypeptides share at least 60%, and preferably 70% amino acid sequence identity with the C-terminal 102-106 amino acids, defining the conserved seven-cysteine domain of human OP-1 and related proteins.

[0037] Amino acid sequence homology can be determined by methods well known in the art. For instance, to determine the percent homology of a candidate amino acid sequence to the sequence of the seven-cysteine domain, the two sequences are first aligned. The alignment can be made with, *e.g.*, the dynamic programming algorithm described in Needleman *et al.*, J. Mol. Biol., 48, pp. 443 (1970), and the Align Program, a commercial software package produced by DNAstar, Inc. The teachings by both sources are incorporated by reference herein. An initial alignment can be refined by comparison to a multi-sequence alignment of a family of related proteins. Once the alignment is made and refined, a percent homology score is calculated. The aligned amino acid residues of the two sequences are compared sequentially for their similarity to each other. Similarity factors include similar size, shape and electrical charge. One particularly preferred method of determining amino acid similarities is the PAM250 matrix described in Dayhoff *et al.*, Atlas of Protein Sequence and Structure, 5, pp. 345-352 (1978 & Supp.), which is incorporated herein by reference. A similarity score is first calculated as the sum of the aligned pair wise amino acid similarity scores.

Insertions and deletions are ignored for the purposes of percent homology and identity. Accordingly, gap penalties are not used in this calculation. The raw score is then normalized by dividing it by the geometric mean of the scores of the candidate sequence and the seven-cysteine domain. The geometric mean is the square root of the product of these scores. The normalized raw score is the percent homology.

[0038] The term "conservative substitutions" refers to residues that are physically or functionally similar to the corresponding reference residues. That is, a conservative substitution and its reference residue have similar size, shape, electric charge, chemical properties including the ability to form covalent or hydrogen bonds, or the like. Preferred conservative substitutions are those fulfilling the criteria defined for an accepted point mutation in Dayhoff *et al., supra*. Examples of conservative substitutions are substitutions within the following groups: (a) valine, glycine; (b) glycine, alanine; (c) valine, isoleucine, leucine; (d) aspartic acid, glutamic acid; (e) asparagine, glutamine; (f) serine, threonine; (g) lysine, arginine, methionine; and (h) phenylalanine, tyrosine. The term "conservative variant" or "conservative variation" also includes the use of a substituting amino acid residue in place of an amino acid residue in a given parent amino acid sequence, where antibodies specific for the parent sequence are also specific for, i.e., "cross-react" or "immuno-react" with, the resulting substituted polypeptide sequence.

[0039] The term "fragment thereof" or "fragment" refers to a stretch of at least about 5 amino acid residues. In some embodiments, this term refers to a

stretch of at least about 10 amino acid residues. In other embodiments, it refers to a stretch of at least about 15 to 20 amino acid residues. The fragments may be naturally derived or synthetically generated. To be active, any fragment must have sufficient length to display biological activity.

Methods Using Smads

[0040] The present invention provides a method of inducing the expression of a Smad in a cell or tissue comprising the step of contacting the cell or tissue capable of expressing the Smad with a bone morphogenic protein.

[0041] This invention also provides gene therapy methods for inducing tissue formation at a target locus, repairing a tissue defect or regenerating tissue at a target locus using Smads. In some embodiments, the methods comprise the step of administering to the target locus a nucleic acid encoding a Smad. In some embodiments, the methods comprise the step of administering to the target locus a vector comprising a nucleic acid encoding a Smad operably linked to an expression control sequence. In some embodiments, the methods comprise the step of administering a cell comprising a vector comprising a nucleic acid encoding a Smad operably linked to an expression control sequence. In some embodiments, the expression control sequence operably linked to a Smad nucleic acid comprises an inducible promoter. In some embodiments, the expression control sequence operably linked to a Smad nucleic acid comprises a constitutive promoter.

[0042] A Smad according to the present invention is a R-Smad. R-Smads includes, but are not limited to,

Smad1, Smad2, Smad3, Smad5, Smad8 or fragments thereof. In a preferred embodiment, the Smad is Smad5. In another embodiment the Smad is recombinant.

[0043] In some embodiments, the gene therapy methods further comprise the step of administering to the target locus a pharmaceutically effective amount of a BMP. In some embodiments, the methods further comprise the step of administering a nucleic acid encoding a BMP. In some embodiments, the methods further comprise the step of administering a vector comprising a nucleic acid encoding a BMP operably linked to an expression control sequence. In some embodiments, the methods further comprise the step of administering a cell comprising a vector comprising a nucleic acid encoding a BMP operably linked to an expression control sequence. In some embodiments, the expression control sequence operably linked to a BMP nucleic acid comprises an inducible promoter. In some embodiments, the expression control sequence operably linked to a BMP nucleic acid comprises a constitutive promoter.

[0044] The BMPs according to the invention include but are not limited to, OP-1 (BMP-7), OP-2, OP-3, COP-1, COP-3, COP-4, COP-5, COP-7, COP-16, BMP-2, BMP-3, BMP-3b, BMP-4, BMP-5, BMP-6, BMP-9, BMP-10, BMP-11, CDMP-3 (BMP-12), CDMP-2 (BMP-13), CDMP-1 (BMP-14), BMP-15, BMP-16, BMP-17, BMP-18, GDF-1, GDF-2, GDF-3, GDF-5, GDF-6, GDF-7, GDF-8, GDF-9, GDF-10, GDF-11, GDF-12, MP121, dorsalin-1, DPP, Vg-1, Vgr-1, 60A protein, NODAL, UNIVIN, SCREW, ADMP, and NEURAL. (see *infra*, for discussion of BMPs). In a preferred embodiment, the BMP is OP-1 (BMP-7). In another preferred embodiment, the BMP is CDMP-1 or GDF-5.

[0045] In some embodiments, more than one BMP is administered. In one preferred embodiment, two BMPs are administered. In yet another preferred embodiment, the first BMP is OP-1 and the second BMP is CDMP-1 or GDF-5. In another embodiment three BMPs are administered. In some embodiments, the BMP is recombinant.

[0046] In some embodiments, the gene therapy methods of this invention further comprise the step of administering a serine/threonine kinase receptor. In some embodiments, the methods of this invention further comprise the step of administering to the target locus a nucleic acid encoding a serine/threonine kinase receptor. In some embodiments, the methods of this invention further comprise the step of administering to the target locus a vector comprising a nucleic acid encoding a serine/threonine kinase receptor operably linked to an expression control sequence. In some embodiments, the methods of this invention further comprise the step of administering to the target locus a cell comprising a vector comprising a nucleic acid encoding a serine/threonine kinase receptor operably linked to an expression control sequence. In some embodiments, the serine/threonine kinase receptor is selected from the group consisting of type I and type II receptors. In some embodiments, only a type I receptor is used. In some embodiments, only a type II receptors is used. In some embodiments both type I and type II receptors are used. Preferably the type I and type II receptors are recombinant. In some embodiments, the expression control sequence operably linked to a serine/threonine kinase receptor nucleic acid comprises an inducible promoter. In some embodiments, the expression control sequence operably linked to a

serine/threonine kinase receptor nucleic acid comprises a constitutive promoter.

[0047] In some embodiments, the type I receptor is an activin receptor-like kinase (ALK). The ALKs according to this invention include, but are not limited to ALK-1, ALK-2, ALK-3, ALK-4, ALK-5, ALK-6, ALK-7, and fragments thereof. Preferred ALKs are ALK-2, ALK-3 and ALK-6.

Bone Morphogenic Protein (BMP) Family

[0048] The BMP family, named for its representative bone morphogenic/osteogenic protein family members, belongs to the TGF-β protein superfamily. Of the reported "BMPs" (BMP-1 to BMP-18), isolated primarily based on sequence homology, all but BMP-1 remain classified as members of the BMP family of morphogenic proteins (Ozkaynak *et al.*, *EMBO J.*, 9, pp. 2085-93 (1990)).

[0049] The BMP family includes other structurally-related members which are bone morphogenic proteins, including the *drosophila* decapentaplegic gene complex (DPP) products, the Vg1 product of *Xenopus laevis* and its murine homolog, Vgr-1 (see, *e.g.*, Massagué, *Annu. Rev. Cell Biol.*, 6, pp. 597-641 (1990), incorporated herein by reference).

[0050] The *Drosophila* DPP and *Xenopus* Vg-1 gene products are 50% identical to each other (and 35-40% identical to TGF-ß). Both the Dpp and Vg-1 products are morphogenic proteins that participate in early patterning events during embryogenesis of their respective hosts. These products appear to be most closely related to mammalian bone morphogenetic proteins BMP-2 and BMP-4, whose C-terminal domains are 75% identical with that of Dpp.

[0051] The C-terminal domains of BMP-3, BMP-5, BMP-6, and OP-1 (BMP-7) are about 60% identical to that of BMP-2, and the C-terminal domains of BMP-6 and OP-1 are 87% identical. BMP-6 is likely the human homolog of the murine Vgr-1 (Lyons *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 86, pp. 4554-59 (1989)); the two proteins are 92% identical overall at the amino acid sequence level (U.S. Patent No. 5,459,047, incorporated herein by reference). BMP-6 is 58% identical to the *Xenopus* Vg-1 product.

[0052] The naturally occurring bone morphogenic proteins share substantial amino acid sequence homology in their C-terminal regions (domains). Typically, the above-mentioned naturally occurring osteogenic proteins are translated as a precursor, having an N-terminal signal peptide sequence typically less than about 30 residues, followed by a "pro" domain that is cleaved to yield the mature C-terminal domain of approximately 100-140 amino acids. The signal peptide is cleaved rapidly upon translation, at a cleavage site that can be predicted in a given sequence using the method of Von Heijne, *Nucleic Acids Research*, 14, pp. 4683-4691 (1986). The pro domain typically is about three times larger than the fully processed mature C-terminal domain.

[0053] Another characteristic of the BMP protein family members is their ability to dimerize. Several bone-derived osteogenic proteins (OPs) and BMPs are found as homo- and heterodimers in their active forms. The ability of OPs and BMPs to form heterodimers may confer additional or altered morphogenic inductive capabilities on bone morphogenic proteins. Heterodimers may exhibit qualitatively or quantitatively different

binding affinities than homodimers for OP and BMP receptor molecules. Altered binding affinities may in turn lead to differential activation of receptors that mediate different signaling pathways, which may ultimately lead to different biological activities or outcomes. Altered binding affinities could also be manifested in a tissue or cell type-specific manner, thereby inducing only particular progenitor cell types to undergo proliferation and/or differentiation.

[0054] In one preferred embodiment of this invention, the BMPs independently comprise a pair of subunits disulfide bonded to produce a dimeric species, wherein at least one of the subunits comprises a recombinant peptide belonging to the BMP protein family. In another preferred embodiment of this invention, the BMPs independently comprise a pair of subunits that produce a dimeric species formed through non-covalent interactions, wherein at least one of the subunits comprises a recombinant peptide belonging to the BMP protein family. Non-covalent interactions include Van der Waals, hydrogen bond, hydrophobic and electrostatic interactions. The dimeric species may be a homodimer or heterodimer and is capable of inducing cell proliferation and/or tissue formation. In some embodiments, the BMPs are each independently monomers.

[0055] In preferred embodiments, the pair of morphogenic polypeptides have amino acid sequences each comprising a sequence that shares a defined relationship with an amino acid sequence of a reference morphogen. Herein, preferred osteogenic polypeptides share a defined relationship with a sequence present in osteogenically active human OP-1, SEQ ID NO: 1. However, any one or more of the naturally occurring or

biosynthetic sequences disclosed herein similarly could be used as a reference sequence. Preferred osteogenic polypeptides share a defined relationship with at least the C-terminal six cysteine domain of human OP-1, residues 335-431 of SEQ ID NO: 1. Preferably, osteogenic polypeptides share a defined relationship with at least the C-terminal seven cysteine domain of human OP-1, residues 330-431 of SEQ ID NO: 1. That is, preferred polypeptides in a dimeric protein with bone morphogenic activity each comprise a sequence that corresponds to a reference sequence or is functionally equivalent thereto.

[0056] Functionally equivalent sequences include functionally equivalent arrangements of cysteine residues disposed within the reference sequence, including amino acid insertions or deletions which alter the linear arrangement of these cysteines, but do not materially impair their relationship in the folded structure of the dimeric morphogen protein, including their ability to form such intra- or inter-chain disulfide bonds as may be necessary for morphogenic activity. Functionally equivalent sequences further include those wherein one or more amino acid residues differs from the corresponding residue of a reference sequence, *e.g.*, the C-terminal seven cysteine domain (also referred to herein as the conserved seven cysteine skeleton) of human OP-1, provided that this difference does not destroy bone morphogenic activity. Accordingly, conservative substitutions of corresponding amino acids in the reference sequence are preferred. Particularly preferred conservative substitutions are those fulfilling the criteria defined for an accepted

point mutation in Dayhoff *et al.*, *supra*, the teachings of which are incorporated by reference herein.

[0057] The osteogenic protein OP-1 has been described (see, *e.g.*, Oppermann *et al.*, U. S. Patent No. 5,354,557, incorporated herein by reference). Natural-sourced osteogenic protein in its mature, native form is a glycosylated dimer typically having an apparent molecular weight of about 30-36 kDa as determined by SDS-PAGE. When reduced, the 30 kDa protein gives rise to two glycosylated peptide subunits having apparent molecular weights of about 16 kDa and 18 kDa. The unglycosylated protein, which also has osteogenic activity, has an apparent molecular weight of about 27 kDa. When reduced, the 27 kDa protein gives rise to two unglycosylated polypeptides, having molecular weights of about 14 kDa to 16 kDa, capable of inducing endochondral bone formation in a mammal. Osteogenic proteins may include forms having varying glycosylation patterns, varying N-termini, and active truncated or mutated forms of native protein.

[0058] As described above, particularly useful sequences include those comprising the C-terminal 96 or 102 amino acid sequences of DPP (from Drosophila), Vg1 (from Xenopus), Vgr-1 (from mouse), the OP-1 and OP-2 proteins,(see U.S. Pat. No. 5,011,691 and Oppermann *et al.*, incorporated herein by reference), as well as the proteins referred to as BMP-2, BMP-3, BMP-4 (see WO 88/00205, U.S. Patent No. 5,013,649 and WO 91/18098, incorporated herein by reference), BMP-5 and BMP-6 (see WO 90/11366, PCT/US90/01630, incorporated herein by reference), BMP-8 and BMP-9.

[0059] Preferred BMPs of this invention comprise at least one polypeptide selected from the group consisting

of OP-1 (BMP-7), OP-2, OP-3, COP-1, COP-3, COP-4, COP-5, COP-7, COP-16, BMP-2, BMP-3, BMP-3b, BMP-4, BMP-5, BMP-6, BMP-9, BMP-10, BMP-11, CDMP-3 (BMP-12), CDMP-2 (BMP-13), CDMP-1 (BMP-14), BMP-15, BMP-16, BMP-17, BMP-18, GDF-1, GDF-2, GDF-3, GDF-5, GDF-6, GDF-7, GDF-8, GDF-9, GDF-10, GDF-11, GDF-12, MP121, dorsalin-1, DPP, Vg-1, Vgr-1, 60A protein, NODAL, UNIVIN, SCREW, ADMP, NEURAL and amino acid sequence variants and homologs thereof, including species homologs thereof and fragments thereof. In some embodiments, one BMP is used. In some embodiments, more than one BMP is used. In some embodiments, two BMPs are used. In some embodiments, three BMPs are used. In some embodiments, the first BMP is OP-1 (BMP-7) or a fragment thereof, and the second BMP is selected from the group consisting of CDMP-1, CDMP-2, CDMP3, GDF-5, GDF-6, GDF-7 and fragments thereof. In some embodiments, the second BMP is CDMP-1, GDF-5 or a fragment thereof. In some embodiments, the first BMP is OP-1 and the second BMP is CDMP-2 or GDF-5.

[0060] Publications disclosing these sequences, as well as their chemical and physical properties, include: OP-1 and OP-2 (U.S. Patent No. 5,011,691; U.S. Patent No. 5,266,683; Ozkaynak *et al.*, *EMBO J.*, 9, pp. 2085-2093 (1990); OP-3 (WO 94/10203 (PCT US93/10520)), BMP-2, BMP-3, BMP-4, (WO 88/00205; Wozney *et al.* *Science*, 242, pp. 1528-1534 (1988)), BMP-5 and BMP-6, (Celeste *et al.*, *PNAS*, 87, 9843-9847 (1991)), Vgr-1 (Lyons *et al.*, *PNAS*, 86, pp. 4554-4558 (1989)); DPP (Padgett *et al.* *Nature*, 325, pp. 81-84 (1987)); Vg-1 (Weeks, *Cell*, 51, pp. 861-867 (1987)); BMP-9 (WO95/33830 (PCT/US95/07084); BMP-10 (WO 94/26893 (PCT/US94/05290); BMP-11 (WO 94/26892 (PCT/US94/05288); BMP-12 (WO95/16035 (PCT/US94/14030); BMP-13 (WO95/16035 (PCT/US94/14030);

GDF-1 (WO 92/00382 (PCT/US91/04096) and Lee *et al. PNAS*, 88, pp. 4250-4254 (1991); GDF-8 (WO 94/21681 (PCT/US94/03019); GDF-9 (WO 94/15966 (PCT/US94/00685); GDF-10 (WO 95/10539 (PCT/US94/11440); GDF-11 (WO 96/01845 (PCT/US95/08543); BMP-15 (WO 96/36710 (PCT/US96/06540); MP-121 (WO 96/01316 (PCT/EP95/02552); GDF-5 (CDMP-1, MP52) (WO 94/15949 (PCT/US94/00657) and WO 96/14335 (PCT/US94/12814) and WO 93/16099 (PCT/EP93/00350)); GDF-6 (CDMP-2, BMP13) (WO 95/01801 (PCT/US94/07762) and WO 96/14335 and WO 95/10635 (PCT/US94/14030)); GDF-7 (CDMP-3, BMP12) (WO 95/10802 (PCT/US94/07799) and WO 95/10635 (PCT/US94/14030)). The above publications are incorporated herein by reference.

[0061] In another embodiment of this invention, the BMPs may be prepared synthetically. BMPs prepared synthetically may be native, or may be non-native proteins, i.e., those not otherwise found in nature. Non-native osteogenic proteins have been synthesized using a series of consensus DNA sequences (U.S. Patent No. 5,324,819, incorporated herein by reference). These consensus sequences were designed based on partial amino acid sequence data obtained from natural osteogenic products and on their observed homologies with other genes reported in the literature having a presumed or demonstrated developmental function.

[0062] Several of the biosynthetic consensus sequences (called consensus osteogenic proteins or "COPs") have been expressed as fusion proteins in prokaryotes. Purified fusion proteins may be cleaved, refolded, implanted in an established animal model and shown to have bone- and/or cartilage-inducing activity. The currently preferred synthetic osteogenic proteins comprise two synthetic amino acid sequences designated

COP-5 (SEQ. ID NO: 2) and COP-7 (SEQ. ID NO: 3). Oppermann *et al.*, U. S. Patent Nos. 5,011,691 and 5,324,819, which are incorporated herein by reference, describe the amino acid sequences of COP-5 and COP-7 as shown below:

COP5 LYVDFS-DVGW**D**DWIVAPPGY**Q**AFYCHGECPFPLAD

COP7 LYVDFS-DVGW**N**DWIVAPPGY**H**AFYCHGECPFPLAD

COP5 H**F**NSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA

COP7 H**L**NSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA

COP5 ISMLYLDENEKVVLKYNQEMVVEGCGCR

COP7 ISMLYLDENEKVVLKYNQEMVVEGCGCR

[0063] In these amino acid sequences, the dashes (-) are used as fillers only to line up comparable sequences in related proteins. Differences between the aligned amino acid sequences are highlighted.

[0064] The DNA and amino acid sequences of these and other BMP family members are published and may be used by those of skill in the art to determine whether a newly identified protein belongs to the BMP family.

[0065] In certain preferred embodiments, the BMPs useful herein independently include those in which the amino acid sequences comprise a sequence sharing at least 70% amino acid sequence homology or "similarity", preferably 80%, more preferably 90%, even more preferably 95%, even more preferably 98% homology or similarity, with a reference bone morphogenic protein selected from the foregoing naturally occurring proteins. Preferably, the reference protein is human OP-1, and the reference sequence thereof is the C-terminal seven cysteine domain present in osteogenically

active forms of human OP-1, residues 330-431 of SEQ ID NO: 1. In some embodiments, the BMP comprises a dimeric protein having an amino acid sequence having at least 70% homology within the C-terminal 102-106 amino acids of human OP-1. In certain embodiments, a polypeptide suspected of being functionally equivalent to a reference BMP polypeptide is aligned therewith using the method of Needleman, *et al.*, *supra*, implemented conveniently by computer programs such as the Align program (DNAstar, Inc.). As noted above, internal gaps and amino acid insertions in the candidate sequence are ignored for purposes of calculating the defined relationship, conventionally expressed as a level of amino acid sequence homology or identity, between the candidate and reference sequences. In one preferred embodiment, the reference sequence is OP-1. In another preferred embodiment, the reference sequence is selected from CDMP-1, CDMP-2 or CDMP-3. Bone morphogenic proteins useful herein accordingly include allelic, phylogenetic counterpart and other variants of the preferred reference sequence, whether naturally-occurring or biosynthetically produced (*e.g.*, including "muteins" or "mutant proteins"), as well as novel members of the general morphogenic family of proteins, including those set forth and identified above. Certain particularly preferred bone morphogenic polypeptides share at least 60% amino acid identity with the preferred reference sequence of human OP-1, still more preferably at least 65% amino acid identity therewith.

[0066] In another embodiment, useful BMPs include those sharing the conserved seven cysteine domain and sharing at least 70% amino acid sequence homology (similarity) within the C-terminal active domain, as

defined herein. In still another embodiment, the BMPs of the invention can be defined as osteogenically active proteins having any one of the generic sequences defined herein, including OPX (SEQ ID NO: 4) and Generic Sequences 7 (SEQ ID NO: 5) and 8 (SEQ ID NO: 6), or Generic Sequences 9 (SEQ ID NO: 7) and 10 (SEQ ID NO: 8).

[0067] The family of bone morphogenic polypeptides useful in the present invention, and members thereof, can be defined by a generic amino acid sequence. For example, Generic Sequence 7 (SEQ ID NO: 5) and Generic Sequence 8 (SEQ ID NO: 6) are 97 and 102 amino acid sequences, respectively, and accommodate the homologies shared among preferred protein family members identified to date, including at least OP-1, OP-2, OP-3, CBMP-2A, CBMP-2B, BMP-3, 60A, DPP, Vg1, BMP-5, BMP-6, Vgr-1, and GDF-1. The amino acid sequences for these proteins are described herein and/or in the art, as summarized above. The generic sequences include both the amino acid identity shared by these sequences in the C-terminal domain, defined by the six and seven cysteine skeletons (Generic Sequences 7 and 8, respectively), as well as alternative residues for the variable positions within the sequence. The generic sequences provide an appropriate cysteine skeleton where inter- or intramolecular disulfide bonds can form, and contain certain critical amino acids likely to influence the tertiary structure of the folded proteins. In addition, the generic sequences allow for an additional cysteine at position 36 (Generic Sequence 7) or position 41 (Generic Sequence 8), thereby encompassing the morphogenically active sequences of OP-2 and OP-3.

Generic Sequence 7

			Leu	Xaa	Xaa	Xaa	Phe	Xaa	Xaa
			1				5		
Xaa	Gly	Trp	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Pro
		10					15		
Xaa	Xaa	Xaa	Xaa	Ala	Xaa	Tyr	Cys	Xaa	Gly
		20					25		
Xaa	Cys	Xaa	Xaa	Pro	Xaa	Xaa	Xaa	Xaa	Xaa
		30					35		
Xaa	Xaa	Xaa	Asn	His	Ala	Xaa	Xaa	Xaa	Xaa
		40					45		
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
		50					55		
Xaa	Xaa	Xaa	Cys	Cys	Xaa	Pro	Xaa	Xaa	Xaa
		60					65		
Xaa	Xaa	Xaa	Xaa	Xaa	Leu	Xaa	Xaa	Xaa	Xaa
		70					75		
Xaa	Xaa	Xaa	Val	Xaa	Leu	Xaa	Xaa	Xaa	Xaa
		80					85		
Xaa	Met	Xaa	Val	Xaa	Xaa	Cys	Xaa	Cys	Xaa
		90					95		

wherein each Xaa independently is selected from a group of one or more specified amino acids defined as follows: "res." means "residue" and Xaa at res.2 = (Tyr or Lys); Xaa at res.3 = Val or Ile); Xaa at res.4 = (Ser, Asp or Glu); Xaa at res.6 = (Arg, Gln, Ser, Lys or Ala); Xaa at res.7 = (Asp or Glu); Xaa at res.8 = (Leu, Val or Ile); Xaa at res. 11 = (Gln, Leu, Asp, His, Asn or Ser); Xaa at res.12 = (Asp, Arg, Asn or Glu); Xaa at res.13 = (Trp or Ser); Xaa at res.14 = (Ile or Val); Xaa at res.15 = (Ile or Val); Xaa at res.16 (Ala or Ser); Xaa at res.18 = (Glu, Gln, Leu, Lys, Pro or Arg); Xaa at res.19 = (Gly or Ser); Xaa at res.20 = (Tyr or Phe); Xaa at res.21 = (Ala, Ser, Asp, Met, His, Gln, Leu or Gly); Xaa at res.23 = (Tyr, Asn or Phe); Xaa at res.26 = (Glu, His, Tyr, Asp, Gln, Ala or Ser); Xaa at res.28 = (Glu, Lys, Asp, Gln or Ala); Xaa at res.30 = (Ala, Ser, Pro, Gln,

Ile or Asn); Xaa at res.31 = (Phe, Leu or Tyr); Xaa at res.33 = (Leu, Val or Met); Xaa at res.34 = (Asn, Asp, Ala, Thr or Pro); Xaa at res.35 = (Ser, Asp, Glu, Leu, Ala or Lys); Xaa at res.36 = (Tyr, Cys, His, Ser or Ile); Xaa at res.37 = (Met, Phe, Gly or Leu); Xaa at res.38 = (Asn, Ser or Lys); Xaa at res.39 = (Ala, Ser, Gly or Pro); Xaa at res.40 = (Thr, Leu or Ser); Xaa at res.44 = (Ile, Val or Thr); Xaa at res.45 = (Val, Leu, Met or Ile); Xaa at res.46 = (Gln or Arg); Xaa at res.47 = (Thr, Ala or Ser); Xaa at res.48 = (Leu or Ile); Xaa at res.49 = (Val or Met); Xaa at res.50 = (His, Asn or Arg); Xaa at res.51 = (Phe, Leu, Asn, Ser, Ala or Val); Xaa at res.52 = (Ile, Met, Asn, Ala, Val, Gly or Leu); Xaa at res.53 = (Asn, Lys, Ala, Glu, Gly or Phe); Xaa at res.54 = (Pro, Ser or Val); Xaa at res.55 = (Glu, Asp, Asn, Gly, Val, Pro or Lys); Xaa at res.56 = (Thr, Ala, Val, Lys, Asp, Tyr, Ser, Gly, Ile or His); Xaa at res.57 = (Val, Ala or Ile); Xaa at res.58 = (Pro or Asp); Xaa at res.59 = (Lys, Leu or Glu); Xaa at res.60 = (Pro, Val or Ala); Xaa at res.63 = (Ala or Val); Xaa at res.65 = (Thr, Ala or Glu); Xaa at res.66 = (Gln, Lys, Arg or Glu); Xaa at res.67 = (Leu, Met or Val); Xaa at res.68 = (Asn, Ser, Asp or Gly); Xaa at res.69 =(Ala, Pro or Ser); Xaa at res.70 = (Ile, Thr, Val or Leu); Xaa at res.71 = (Ser, Ala or Pro); Xaa at res.72 = (Val, Leu, Met or Ile); Xaa at res.74 = (Tyr or Phe); Xaa at res.75 = (Phe, Tyr, Leu or His); Xaa at res.76 = (Asp, Asn or Leu); Xaa at res.77 = (Asp, Glu, Asn, Arg or Ser); Xaa at res.78 = (Ser, Gln, Asn, Tyr or Asp); Xaa at res.79 = (Ser, Asn, Asp, Glu or Lys); Xaa at res.80 = (Asn, Thr or Lys); Xaa at res.82 = (Ile, Val or Asn); Xaa at res.84 = (Lys or Arg); Xaa at res.85 = (Lys, Asn, Gln, His, Arg or Val); Xaa at res.86 = (Tyr, Glu or His); Xaa

at res.87 = (Arg, Gln, Glu or Pro); Xaa at res.88 = (Asn, Glu, Trp or Asp); Xaa at res.90 = (Val, Thr, Ala or Ile); Xaa at res.92 = (Arg, Lys, Val, Asp, Gln or Glu); Xaa at res.93 = (Ala, Gly, Glu or Ser); Xaa at res.95 = (Gly or Ala) and Xaa at res.97 = (His or Arg).

[0068] Generic Sequence 8 (SEQ ID NO: 6) includes all of Generic Sequence 7 and in addition includes the following sequence (SEQ ID NO: 9) at its N-terminus:

SEQ ID NO: 9

Cys	Xaa	Xaa	Xaa	Xaa
1				5

Accordingly, beginning with residue 7, each "Xaa" in Generic Sequence 8 is a specified amino acid defined as for Generic Sequence 7, with the distinction that each residue number described for Generic Sequence 7 is shifted by five in Generic Sequence 8. Thus, "Xaa at res.2 =(Tyr or Lys)" in Generic Sequence 7 refers to Xaa at res.7 in Generic Sequence 8. In Generic Sequence 8, Xaa at res.2 = (Lys, Arg, Ala or Gln); Xaa at res.3 = (Lys, Arg or Met); Xaa at res.4 = (His, Arg or Gln); and Xaa at res.5 = (Glu, Ser, His, Gly, Arg, Pro, Thr, or Tyr).

[0069] In another embodiment, useful osteogenic proteins include those defined by Generic Sequences 9 and 10, defined as follows.

[0070] Specifically, Generic Sequences 9 and 10 are composite amino acid sequences of the following proteins: human OP-1, human OP-2, human OP-3, human BMP-2, human BMP-3, human BMP-4, human BMP-5, human BMP-6, human BMP-8, human BMP-9, human BMP-10, human BMP-11, Drosophila 60A, Xenopus Vg-1, sea urchin UNIVIN, human CDMP-1 (mouse GDF-5), human CDMP-2 (mouse GDF-6, human

BMP-13), human CDMP-3 (mouse GDF-7, human BMP-12), mouse GDF-3, human GDF-1, mouse GDF-1, chicken DORSALIN, dpp, Drosophila SCREW, mouse NODAL, mouse GDF-8, human GDF-8, mouse GDF-9, mouse GDF-10, human GDF-11, mouse GDF-11, human BMP-15, and rat BMP3b. Like Generic Sequence 7, Generic Sequence 9 is a 97 amino acid sequence that accommodates the C-terminal six cysteine skeleton and, like Generic Sequence 8, Generic Sequence 10 is a 102 amino acid sequence which accommodates the seven cysteine skeleton.

Generic Sequence 9

Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
1				5					10
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Pro	Xaa	Xaa	Xaa
				15					20
Xaa	Xaa	Xaa	Xaa	Cys	Xaa	Gly	Xaa	Cys	Xaa
				25					30
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
				35					40
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
				45					50
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
				55					60
Xaa	Cys	Xaa	Pro	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
				65					70
Xaa	Xaa	Leu	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
				75					80
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
				85					90
Xaa	Xaa	Xaa	Cys	Xaa	Cys	Xaa			
				95					

wherein each Xaa is independently selected from a group of one or more specified amino acids defined as follows: "res." means "residue" and Xaa at res.1 = (Phe, Leu or Glu); Xaa at res.2 = (Tyr, Phe, His, Arg, Thr, Lys, Gln, Val or Glu); Xaa at res.3 = (Val, Ile, Leu or Asp); Xaa at res.4 = (Ser, Asp, Glu, Asn or Phe); Xaa at res.5 = (Phe or Glu); Xaa at res.6 = (Arg, Gln, Lys, Ser, Glu,

Ala or Asn); Xaa at res.7 = (Asp, Glu, Leu, Ala or Gln); Xaa at res.8 = (Leu, Val, Met, Ile or Phe); Xaa at res.9 = (Gly, His or Lys); Xaa at res.10 = (Trp or Met); Xaa at res.11 = (Gln, Leu, His, Glu, Asn, Asp, Ser or Gly); Xaa at res.12 = (Asp, Asn, Ser, Lys, Arg, Glu or His); Xaa at res.13 = (Trp or Ser); Xaa at res.14 = (Ile or Val); Xaa at res.15 = (Ile or Val); Xaa at res.16 = (Ala, Ser, Tyr or Trp); Xaa at res. 18 = (Glu, Lys, Gln, Met, Pro, Leu, Arg, His or Lys); Xaa at res.19 = (Gly, Glu, Asp, Lys, Ser, Gln, Arg or Phe); Xaa at res.20 = (Tyr or Phe); Xaa at res.21 = (Ala, Ser, Gly, Met, Gln, His, Glu, Asp, Leu, Asn, Lys or Thr); Xaa at res.22 = (Ala or Pro); Xaa at res.23 = (Tyr, Phe, Asn, Ala or Arg); Xaa at res.24 = (Tyr, His, Glu, Phe or Arg); Xaa at res.26 = (Glu, Asp, Ala, Ser, Tyr, His, Lys, Arg, Gln or Gly); Xaa at res.28 = (Glu, Asp, Leu, Val, Lys, Gly, Thr, Ala or Gln); Xaa at res.30 = (Ala, Ser, Ile, Asn, Pro, Glu, Asp, Phe, Gln or Leu); Xaa at res.31= (Phe, Tyr, Leu, Asn, Gly or Arg); Xaa at res.32 = (Pro, Ser, Ala or Val); Xaa at res.33 = (Leu, Met, Glu, Phe or Val); Xaa at res.34 = (Asn, Asp, Thr, Gly, Ala, Arg, Leu or Pro); Xaa at res.35 = (Ser, Ala, Glu, Asp, Thr, Leu, Lys, Gln or His); Xaa at res.36 = (Tyr, His, Cys, Ile, Arg, Asp, Asn, Lys, Ser, Glu or Gly); Xaa at res.37 = (Met, Leu, Phe, Val, Gly or Tyr); Xaa at res.38 = (Asn, Glu, Thr, Pro, Lys, His, Gly, Met, Val or Arg); Xaa at res.39 = (Ala, Ser, Gly, Pro or Phe); Xaa at res.40 = (Thr, Ser, Leu, Pro, His or Met); Xaa at res.41 = (Asn, Lys, Val, Thr or Gln); Xaa at res.42 = (His, Tyr or Lys); Xaa at res.43 = (Ala, Thr, Leu or Tyr); Xaa at res.44 = (Ile, Thr, Val, Phe, Tyr, Met or Pro); Xaa at res.45 = (Val, Leu, Met, Ile or His); Xaa at res.46 = (Gln, Arg or Thr); Xaa at res.47 = (Thr, Ser, Ala, Asn

or His); Xaa at res.48 = (Leu, Asn or Ile); Xaa at res.49 = (Val, Met, Leu, Pro or Ile); Xaa at res.50 = (His, Asn, Arg, Lys, Tyr or Gln); Xaa at res.51 = (Phe, Leu, Ser, Asn, Met, Ala, Arg, Glu, Gly or Gln); Xaa at res.52 = (Ile, Met, Leu, Val, Lys, Gln, Ala or Tyr); Xaa at res.53 = (Asn, Phe, Lys, Glu, Asp, Ala, Gln, Gly, Leu or Val); Xaa at res.54 = (Pro, Asn, Ser, Val or Asp); Xaa at res.55 = (Glu, Asp, Asn, Lys, Arg, Ser, Gly, Thr, Gln, Pro or His); Xaa at res.56 = (Thr, His, Tyr, Ala, Ile, Lys, Asp, Ser, Gly or Arg); Xaa at res.57 = (Val, Ile, Thr, Ala, Leu or Ser); Xaa at res.58 = (Pro, Gly, Ser, Asp or Ala); Xaa at res.59 = (Lys, Leu, Pro, Ala, Ser, Glu, Arg or Gly); Xaa at res.60 = (Pro, Ala, Val, Thr or Ser); Xaa at res.61 = (Cys, Val or Ser); Xaa at res.63 = (Ala, Val or Thr); Xaa at res.65 = (Thr, Ala, Glu, Val, Gly, Asp or Tyr); Xaa at res.66 = (Gln, Lys, Glu, Arg or Val); Xaa at res.67 = (Leu, Met, Thr or Tyr); Xaa at res.68 = (Asn, Ser, Gly, Thr, Asp, Glu, Lys or Val); Xaa at res.69 = (Ala, Pro, Gly or Ser); Xaa at res.70 = (Ile, Thr, Leu or Val); Xaa at res.71 = (Ser, Pro, Ala, Thr, Asn or Gly); Xaa at res.72 = (Val, Ile, Leu or Met); Xaa at res.74 = (Tyr, Phe, Arg, Thr, Tyr or Met); Xaa at res.75 = (Phe, Tyr, His, Leu, Ile, Lys, Gln or Val); Xaa at res.76 = (Asp, Leu, Asn or Glu); Xaa at res.77 = (Asp, Ser, Arg, Asn, Glu, Ala, Lys, Gly or Pro); Xaa at res.78 = (Ser, Asn, Asp, Tyr, Ala, Gly, Gln, Met, Glu, Asn or Lys); Xaa at res.79 = (Ser, Asn, Glu, Asp, Val, Lys, Gly, Gln or Arg); Xaa at res.80 = (Asn, Lys, Thr, Pro, Val, Ile, Arg, Ser or Gln); Xaa at res.81 = (Val, Ile, Thr or Ala); Xaa at res.82 = (Ile, Asn, Val, Leu, Tyr, Asp or Ala); Xaa at res.83 = (Leu, Tyr, Lys or Ile); Xaa at res.84 = (Lys, Arg, Asn, Tyr, Phe, Thr, Glu or Gly); Xaa at res.85 = (Lys, Arg, His,

Gln, Asn, Glu or Val); Xaa at res.86 = (Tyr, His, Glu or Ile); Xaa at res.87 = (Arg, Glu, Gln, Pro or Lys); Xaa at res.88 = (Asn, Asp, Ala, Glu, Gly or Lys); Xaa at res.89 = (Met or Ala); Xaa at res.90 = (Val, Ile, Ala, Thr, Ser or Lys); Xaa at res.91 = (Val or Ala); Xaa at res.92 = (Arg, Lys, Gln, Asp, Glu, Val, Ala, Ser or Thr); Xaa at res.93 = (Ala, Ser, Glu, Gly, Arg or Thr); Xaa at res.95 = (Gly, Ala or Thr); Xaa at res.97 = (His, Arg, Gly, Leu or Ser). Further, after res.53 in rBMP3b and mGDF-10 there is an Ile; after res.54 in GDF-1 there is a T; after res.54 in BMP3 there is a V; after res.78 in BMP-8 and Dorsalin there is a G; after res.37 in hGDF-1 there is Pro, Gly, Gly, Pro.

[0071] Generic Sequence 10 (SEQ ID NO: 8) includes all of Generic Sequence 9 (SEQ ID NO: 7) and in addition includes the following sequence (SEQ ID NO: 9) at its N-terminus:

SEQ ID NO: 9

Cys	Xaa	Xaa	Xaa	Xaa
1				5

Accordingly, beginning with residue 6, each "Xaa" in Generic Sequence 10 is a specified amino acid defined as for Generic Sequence 9, with the distinction that each residue number described for Generic Sequence 9 is shifted by five in Generic Sequence 10. Thus, "Xaa at res.1 = (Tyr, Phe, His, Arg, Thr, Lys, Gln, Val or Glu)" in Generic Sequence 9 refers to Xaa at res.6 in Generic Sequence 10. In Generic Sequence 10, Xaa at res.2 = (Lys, Arg, Gln, Ser, His, Glu, Ala, or Cys); Xaa at res.3 = (Lys, Arg, Met, Lys, Thr, Leu, Tyr, or Ala); Xaa

at res.4 = (His, Gln, Arg, Lys, Thr, Leu, Val, Pro, or Tyr); and Xaa at res.5 = (Gln, Thr, His, Arg, Pro, Ser, Ala, Gln, Asn, Tyr, Lys, Asp, or Leu).

[0072] As noted above, certain currently preferred bone morphogenic polypeptide sequences useful in this invention have greater than 60% identity, preferably greater than 65% identity, with the amino acid sequence defining the preferred reference sequence of hOP-1. These particularly preferred sequences include allelic and phylogenetic counterpart variants of the OP-1 and OP-2 proteins, including the Drosophila 60A protein. Accordingly, in certain particularly preferred embodiments, useful BMPs include active proteins comprising pairs of polypeptide chains within the generic amino acid sequence herein referred to as "OPX" (SEQ ID NO: 4), which defines the seven cysteine skeleton and accommodates the homologies between several identified variants of OP-1 and OP-2. As described therein, each Xaa at a given position independently is selected from the residues occurring at the corresponding position in the C-terminal sequence of mouse or human OP-1 or OP-2.

SEQ ID NO: 4

```
Cys Xaa Xaa His Glu Leu Tyr Val Ser Phe Xaa Asp Leu Gly Trp Xaa Asp Trp
1               5                  10                  15
Xaa Ile Ala Pro Xaa Gly Tyr Xaa Ala Tyr Tyr Cys Glu Gly Glu Cys Xaa Phe Pro
            20                  25                  30                  35
Leu Xaa Ser Xaa Met Asn Ala Thr Asn His Ala Ile Xaa Gln Xaa Leu Val His Xaa
        40                  45                  50                  55
Xaa Xaa Pro Xaa Xaa Val Pro Lys Xaa Cys Cys Ala Pro Thr Xaa Leu Xaa Ala
            60                  65                  70
Xaa Ser Val Leu Tyr Xaa Asp Xaa Ser Xaa Asn Val Ile Leu Xaa Lys Xaa Arg
75                  80                  85                  90
Asn Met Val Val Xaa Ala Cys Gly Cys His
        95                  100
```

wherein Xaa at res.2 = (Lys or Arg); Xaa at res.3 = (Lys or Arg); Xaa at res.11 = (Arg or Gln); Xaa at res.16 = (Gln or Leu); Xaa at res.19 = (Ile or Val); Xaa at res.23 = (Glu or Gln); Xaa at res.26 = (Ala or Ser); Xaa at res.35 = (Ala or Ser); Xaa at res.39 = (Asn or Asp); Xaa at res.41 = (Tyr or Cys); Xaa at res.50 = (Val or Leu); Xaa at res.52 = (Ser or Thr); Xaa at res.56 = (Phe or Leu); Xaa at res.57 = (Ile or Met); Xaa at res.58 = (Asn or Lys); Xaa at res.60 = (Glu, Asp or Asn); Xaa at res.61 = (Thr, Ala or Val); Xaa at res.65 = (Pro or Ala); Xaa at res.71 = (Gln or Lys); Xaa at res.73 = (Asn or Ser); Xaa at res.75 = (Ile or Thr); Xaa at res.80 = (Phe or Tyr); Xaa at res.82 = (Asp or Ser); Xaa at res.84 = (Ser or Asn); Xaa at res.89 = (Lys or Arg); Xaa at res.91 = (Tyr or His); and Xaa at res.97 = (Arg or Lys).

[0073] In still another preferred embodiment, useful BMPs have polypeptide chains with amino acid sequences comprising a sequence encoded by a nucleic acid that hybridizes, under low, medium or high stringency hybridization conditions, to DNA or RNA encoding reference BMP sequences, *e.g.*, C-terminal sequences defining the conserved seven cysteine domains of OP-1, OP-2, BMP-2, BMP-4, BMP-5, BMP-6, 60A, GDF-5, GDF-6, GDF-7 and the like. As used herein, high stringent hybridization conditions are defined as hybridization according to known techniques in 40% formamide, 5 X SSPE, 5 X Denhardt's Solution, and 0.1% SDS at 37°C overnight, and washing in 0.1 X SSPE, 0.1% SDS at 50°C. Standard stringent conditions are well characterized in commercially available, standard molecular cloning texts. See, for example, *Molecular Cloning A Laboratory*

Manual, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volumes I and II (D.N. Glover ed., 1985); *Oligonucleotide Synthesis* (M.J. Gait ed., 1984): *Nucleic Acid Hybridization* (B. D. Hames & S.J. Higgins eds. 1984); and B. Perbal, *A Practical Guide To Molecular Cloning* (1984), the disclosures of which are incorporated herein by reference.

[0074] As noted above, proteins useful in the present invention generally are dimeric proteins comprising a folded pair of the above polypeptides. In some embodiments, the pair of polypeptides are not disulfide bonded. In some embodiments the pair of polypeptides are disulfide bonded. Such disulfide bonded BMPs are inactive when reduced, but are active as oxidized homodimers and when oxidized in combination with others of this invention to produce heterodimers. Thus, members of a folded pair of bone morphogenic polypeptides in a morphogenically active protein can be selected independently from any of the specific polypeptides mentioned above.

[0075] The BMPs useful in the materials and methods of this invention include proteins comprising any of the polypeptide chains described above, whether isolated from naturally-occurring sources, or produced by recombinant DNA or other synthetic techniques, and includes allelic and phylogenetic counterpart variants of these proteins, as well as muteins thereof, and various truncated and fusion constructs. Deletion or addition mutants also are envisioned to be active, including those which may alter the conserved C-terminal six or seven cysteine domain, provided that the alteration does not functionally disrupt the

relationship of these cysteines in the folded structure. Accordingly, such active forms are considered the equivalent of the specifically described constructs disclosed herein. The proteins may include forms having varying glycosylation patterns, varying N-termini, a family of related proteins having regions of amino acid sequence homology, and active truncated or mutated forms of native or biosynthetic proteins, produced by expression of recombinant DNA in host cells.

[0076] The BMPs contemplated herein can be expressed from intact or truncated cDNA or from synthetic DNAs in prokaryotic or eukaryotic host cells, and purified, cleaved, refolded, and dimerized to form morphogenically active compositions. Alternatively, cells expressing recombinant may be used in the methods of this invention. Currently preferred host cells include, without limitation, prokaryotes including *E. coli* or eukaryotes including yeast, or mammalian cells, such as CHO, COS or BSC cells. One of ordinary skill in the art will appreciate that other host cells can be used to advantage. Detailed descriptions of the bone morphogenic proteins useful in the practice of this invention, including how to make, use and test them for osteogenic activity, are disclosed in numerous publications, including U.S. Patent Nos. 5,266,683 and 5,011,691, the disclosures of which are incorporated by reference herein.

[0077] Thus, in view of this disclosure and the knowledge available in the art, skilled genetic engineers can isolate genes from cDNA or genomic libraries of various different biological species, which encode appropriate amino acid sequences, or construct DNAs from oligonucleotides, and then can express them in

various types of host cells, including both prokaryotes and eukaryotes, to produce large quantities of active proteins.

Serine/Threonine Kinase Receptors

[0078] TGF-β superfamily members elicit their cellular responses through formation of heteromeric complexes of specific type I and type II serine/threonine kinase receptors. To date five type II receptors and seven type I receptors (also termed activin receptor-like kinases (ALKs) have been identified (See e.g., Derynck et al., Biochem. Biophys. Acta, 1333, pp. F105-F150 (1997); Massague, Annu. Rev. Biochem., 67, pp. 753-791 (1998)). The type II receptors are kinases and are constitutively active. Upon ligand-mediated heteromeric complex formation, the type II receptor phosphorylates particular serine and threonine residues in the type I receptor juxtamembrane region, thereby activating the type I receptor.

[0079] ALK-1 was identified as an endothelial specific TGF-β type I receptor (see e.g., Oh et al., Proc. Natl. Acad. Sci. U.S.A., 97, pp. 2626-2631 (2000); see also, Lux et al., J. Biol. Chem., 274, 9984-9992 (1999); US patent 5,968,752). ALK-2 was identified as a type I receptor for activin, TGF-β and BMPs (see, e.g., Attisano et al., Cell, 75, pp. 671-680 (1993); Miettinen et al., J. Cell Biol., 127, pp. 2021-2036 (1994); ten Dijke et al., Science 264, 101-104 (1994); Macias-Silva et al., J. Cell Biol., 273, pp. 25628-25636 (1998); US Patent 6,271,365). ALK-4 and ALK-5 were identified as activin and TGF-β type I receptors and ALK-3 and ALK-6 were identified as BMP type I receptors (see, e.g., Oh et al., *supra* and Lux et al., *supra;* US Patents 6,271,365 and 6,207,814). ALK-7 was identified as the

type I receptor for Nodal (see, e.g., Reissman et al., Genes Dev., 15, pp. 2010-22 (2001); US patent 5,891,638).

[0080] In some embodiments of the invention, type II receptors are used. In some embodiments of the invention, type I receptors are used. In yet other embodiments of the invention, both type I and type II receptors are used. In some embodiments, the type I receptor is selected from the group consisting of ALK-1, ALK-2, ALK-3, ALK-4, ALK-5, ALK-6, ALK-7, amino acid sequence variants and homologs thereof, including species homologs thereof and fragments thereof. In a preferred embodiment, the type I receptor is ALK-2, ALK-3 or ALK-6. In some embodiments, the type I and type II receptors are recombinant.

[0081] Publications disclosing these sequences as well as their chemical and physical properties include: ALK-1 (US patent 5,968,752); ALK-2, ALK-4 and ALK-5 (US Patent 6,271,365); ALK-3 and ALK-6 (US Patent 6,207,814); ALK-7 (US patents 5,891,638, 5,614,609 and 5,789,565). These publications are incorporated herein by reference.

[0082] The type I and type II receptors contemplated herein can be expressed from intact or truncated cDNA or from synthetic DNAs in prokaryotic or eukaryotic host cells (see discussion of protein expression, *infra*).

[0083] Thus, in view of this disclosure and the knowledge available in the art, skilled genetic engineers can express these receptors in various types of host cells, including both prokaryotes and eukaryotes.

Smads

[0084] The Smad family of proteins can be divided into three subfamilies: R-Smads, Co-Smads and I-Smads. The subfamily of R-Smads can be divided into two groups: BMP-Smads and TGF-β/activin-Smads. Smad1, Smad5, and Smad8 are phosphorylated by ALK-1, ALK-2, ALK-3 and ALK-6, and Smad2 and Smad3 are activated by ALK-5 and ALK-4, respectively. Smad2 and Smad3 are also activated by ALK-7 (see e.g., Derynck et al., Biochim. Biophys. Acta, 1333, pp. F105-F150 (1997), Massague, Annu. Rev. Biochem., 67, 753-792 (1998); Itoh et al., Eur. J. Biochem., 267, 6954-6967 (2000)). The phosporylated R-Smads form heteromeric complexes with Co-Smads. To date only one Co-Smad has been identified: Smad4. I-Smads, i.e., Smad6 and Smad7 inhibit TGF-β family signaling by preventing the activation of R-Smads and Co-Smads (see, e.g., Itoh et al., *supra*).

[0085] In some embodiments, the Smads used in the methods of this invention are selected from the group consisting of Smad1, Smad2, Smad3, Smad5, Smad8, amino acid sequence variants and homologs thereof, including species homologs thereof and fragments thereof. In a preferred embodiment, the Smad is Smad5. In another embodiments, the Smad is recombinant.

[0086] The sequences for Smads are included in the following publications or are readily available to the skilled worker from GenBank: Smad1 (rat: GenBank Accession No. AF067727; SEQ ID NO: 12 and 13); Smad2 (mouse: GenBank Accession No. NM_010754; SEQ ID NO: 14 and 15), Smad3 (human: GenBank Accession No. NM_005902; SEQ ID NO: 16 and 17), Smad4 (mouse: Genbank Accession No. NM_008540; human: GenBank Accession No. NM_005359; SEQ ID NO: 18 and 19), Smad5 (rat: GenBank Accession No.

NM_021692; SEQ ID NO: 20 and 21), Smad6 (US patents 6,534,476 and 6,270,994; human: GenBank Accession No. NM_005585; SEQ ID NO: 22 and 23), Smad7 (US patents 6,020,464 and 6,251,628; human: GenBank Accession No. NM_005904; SEQ ID NO: 24 and 25), and Smad8 (rat: GenBank Accession No. AF012347; SEQ ID NO: 26 and 27). These publications and GenBank Accesion numbers are incorporated herein by reference.

[0087] The Smads contemplated herein can be expressed from intact or truncated cDNA or from synthetic DNAs in prokaryotic or eukaryotic host cells (see discussion of protein expression, *infra*).

[0088] Thus, in view of this disclosure and the knowledge available in the art, skilled genetic engineers can express these receptors in various types of host cells, including both prokaryotes and eukaryotes (*e.g.*, progenitor cells).

Production or expression of BMPs, serine/threonine kinase receptors and Smads

[0089] The BMPs, serine/threonine kinase receptors and Smads according to this invention may be derived from a variety of sources. They may be isolated from natural sources, or may be produced by expressing an appropriate recombinant DNA molecule in a host cell. In addition, the BMPs, serine/threonine kinase receptors and Smads of this invention may be derived synthetically and synthetic proteins may optionally be expressed from a recombinant DNA molecule in a host cell.

1. Naturally-derived proteins

[0090] In one embodiment of this invention, the BMPs, serine/threonine kinase receptors and Smads are isolated from natural sources. BMPs, serine/threonine kinase

receptors and Smads may be purified from tissue sources, preferably mammalian tissue sources, using conventional physical and chemical separation techniques well known to those of skill in the art.

2. Recombinantly-expressed proteins

[0091] In another embodiment of this invention, the Smads, BMPs and serine/threonine kinase receptors and are produced by the expression of an appropriate recombinant DNA molecule in a host cell. The DNA and amino acid sequences of Smads, BMPs and serine/threonine kinase receptors have been reported, and methods for their recombinant production are published and otherwise known to those of skill in the art. For a general discussion of cloning and recombinant DNA technology, see Ausubel et al., *supra*; see also Watson et al., *Recombinant DNA*, 2d ed. 1992 (W.H. Freeman and Co., New York).

[0092] For cloning and expressing Smads, BMPs and serine/threonine kinase receptors, standard recombinant DNA techniques may be used. With the DNA sequence available, a DNA fragment encoding any of these proteins be inserted into an expression vector selected to work in conjunction with a desired host expression system. The DNA fragment is cloned into the vector with the proper transcription control elements. In some embodiments, the expression of the desired protein may be constitutive. In some embodiments, the expression of the desired protein is under the control of an inducible promoter.

Vectors

[0093] In some embodiments, the invention provides vectors comprising the nucleic acids encoding Smads, serine/threonine kinase receptors and/or BMPs. The

choice of vector and expression control sequences to which the nucleic acids of this invention are operably linked depends on the functional properties desired, *e.g.*, protein expression, and the host cell to be transformed.

[0094] Expression control elements useful for regulating the expression of an operably linked coding sequence are known in the art. Examples include, but are not limited to, inducible promoters, constitutive promoters, secretion signals, and other regulatory elements. When an inducible promoter is used, it can be controlled, e.g., by a change in nutrient status (e.g. concentration of growth factors or BMPs), or a change in temperature, in the host cell medium.

[0095] An appropriate vector is selected according to the host system selected. Useful vectors include but are not limited to plasmids, cosmids, bacteriophage, insect and animal viral vectors, including retroviruses, and other single and double-stranded DNA viruses.

[0096] In some embodiments, it may be preferable to recombinantly produce a mammalian protein for therapeutic uses in mammalian cell culture systems in order to produce a protein whose structure resembles more closely that of the natural material. Recombinant protein production in mammalian cells requires the establishment of appropriate cells and cell lines that are easy to transfect, are capable of stably maintaining foreign DNA with an unrearranged sequence, and which have the necessary cellular components for efficient transcription, translation, post-translational modification and secretion of the protein. In addition, a suitable vector carrying the gene of interest is necessary.

[0097] DNA vector design for transfection into mammalian cells should include appropriate sequences to promote expression of the gene of interest, including: appropriate transcription initiation, termination and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (i.e., Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance protein secretion.

[0098] Preferred DNA vectors also include a marker gene and means for amplifying the copy number of the gene of interest. DNA vectors may also comprise stabilizing sequences (e.g., ori- or ARS-like sequences and telomere-like sequences), or may alternatively be designed to favor directed or non-directed integration into the host cell genome.

[0099] Substantial progress in the development of mammalian cell expression systems has been made in the last decade and many aspects of the system are well characterized. A detailed review of the production of foreign proteins in mammalian cells, including useful cells, protein expression-promoting sequences, marker genes, and gene amplification methods, is disclosed in M. M. Bendig, *Genetic Engineering*, 7, pp. 91-127 (1988).

[0100] Particular details of the transfection, expression and purification of recombinant proteins are well documented and are understood by those of skill in the art. Further details on the various technical aspects of each of the steps used in recombinant production of foreign genes in mammalian cell expression systems can be found in a number of texts and laboratory manuals in the art. See, e.g., F. M. Ausubel et al.,

ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York (1989).

[0101] Briefly, among the best characterized transcription promoters useful for expressing a foreign gene in a particular mammalian cell are the SV40 early promoter, the adenovirus major late promoter (AdMLP), the mouse metallothionein-I promoter (mMT-I), the Rous sarcoma virus (RSV) long terminal repeat (LTR), the mouse mammary tumor virus long terminal repeat (MMTV-LTR), and the human cytomegalovirus major intermediate-early promoter (hCMV). The DNA sequences for all of these promoters are known in the art and are available commercially.

[0102] One method of gene amplification in mammalian cell systems is the use of the selectable dihydrofolate reductase (DHFR) gene in a dhfr- cell line. Generally, the DHFR gene is provided on the vector carrying the gene of interest, and addition of increasing concentrations of the cytotoxic drug methotrexate (MTX) leads to amplification of the DHFR gene copy number, as well as that of the physically-associated gene of interest. DHFR as a selectable, amplifiable marker gene in transfected chinese hamster ovary cell lines (CHO cells) is particularly well characterized in the art. Other useful amplifiable marker genes include the adenosine deaminase (ADA) and glutamine synthetase (GS) genes.

[0103] In one expression system, gene amplification is further enhanced by modifying marker gene expression regulatory sequences (e.g., enhancer, promoter, and transcription or translation initiation sequences) to reduce the levels of marker protein produced. Lowering the level of DHFR transcription increases the DHFR gene

copy number (and the physically-associated gene) to enable the transfected cell to adapt to growth in even low levels of methotrexate (e.g., 0.1 μM MTX). As will be appreciated by those skilled in the art, other useful weak promoters, different from those disclosed and preferred herein, can be constructed using standard vector construction methodologies. In addition, other, different regulatory sequences also can be modified to achieve the same effect.

[0104] Another gene amplification scheme relies on the temperature sensitivity (ts) of BSC40-tsA58 cells transfected with an SV40 vector. Temperature reduction to 33 °C stabilizes the temperature sensitive SV40 T antigen, which leads to the excision and amplification of the integrated transfected vector DNA thereby amplifying the physically associated gene of interest.

[0105] Eukaryotic cell expression vectors are known in the art and are commercially available. Typically, such vectors contain convenient restriction sites for insertion of the desired DNA segment.

[0106] Eukaryotic cell expression vectors may include a selectable marker, e.g., a drug resistance gene. The neomycin phosphotransferase (neo) gene (Southern et al., 1982, J. Mol. Anal. Genet. 1:327-341) is an example of such a gene.

[0107] To express the desired proteins of this invention, DNAs encoding the proteins (Smads, serine/threonine kinase receptors and/or BMPs are inserted into expression vectors such as plasmids, retroviruses, cosmids, YACs, EBV-derived episomes, and the like. The expression vector and expression control sequences are chosen to be compatible with the expression host cell used. In some embodiments, Smad,

serine/threonine kinase receptors and BMP nucleic acids can be inserted into separate vectors. In some embodiments, the Smad, serine/threonine kinase receptor and BMP nucleic acids are inserted into the same expression vector. In some embodiments the Smad and serine/threonine kinase receptor nucleic acids are inserted into the same vector. In some embodiments, the Smad and BMP nucleic acids are inserted into the same vector. Alternatively, any combination of Smad, BMP and/or serine/threonine kinase receptor are inserted into the same vector.

[0108] A convenient vector is one that encodes a functionally complete protein. To the extent secretion of a desired protein is required, the recombinant expression vector can also encode a signal peptide that facilitates secretion of the desired protein from a host cell.

[0109] Nucleic acid molecules encoding Smads, serine/threonine kinase receptors and/or BMPs, and vectors comprising these nucleic acid molecules, can be used for transformation of a suitable host cell. Transformation can be by any suitable method. Methods for introduction of exogenous DNA into mammalian cells are well known in the art and include dextran-mediated transfection, calcium phosphate precipitation, polybrene-mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei. In addition, nucleic acid molecules may be introduced into mammalian cells by viral vectors.

[0110] Transformation of host cells can be accomplished by conventional methods suited to the vector and host cell employed. For transformation of

prokaryotic host cells, electroporation and salt treatment methods can be employed (Cohen et al., 1972, Proc. Natl. Acad. Sci. USA 69:2110-2114). For transformation of vertebrate cells, electroporation, cationic lipid or salt treatment methods can be employed. See, e.g., Graham et al., 1973, Virology 52:456-467; Wigler et al., 1979, Proc. Natl. Acad. Sci. USA 76:1373-1376f.

Host Cells

[0111] Host cells can be prokaryotic or eukaryotic. Useful host cells include but are not limited to bacteria such as *E. coli*, yeasts such as *Saccharomyces* and *Picia*, insect-baculovirus cell system, and primary, transformed or immortalized eukaryotic cells in culture. Preferred eukaryotic host cells include, but are not limited to, yeast and mammalian cells, e.g., Chinese hamster ovary (CHO) cell, NIH Swiss mouse embryo cells NIH-3T3, baby hamster kidney cells (BHK), C2C12 cells and BSC cells. Other useful eukaryotic cells include osteoprogenitor cells, cartilage progenitor cells, tendon progenitor cells, ligament progenitor cells and neural progenitor cells.

[0112] The methodology disclosed herein includes the use of COS cells for the rapid evaluation of vector construction and gene expression, and the use of established cell lines for long term protein production.

[0113] The choice of cells/cell lines is also important and depends on the needs of the skilled practitioner. Monkey kidney cells (COS) provide high levels of transient gene expression providing a useful means for rapidly testing vector construction and the expression of cloned genes. COS cells are transfected

with a simian virus 40 (SV40) vector carrying the gene of interest. The transfected COS cells eventually die, thus preventing the long term production of the desired protein product. However, transient expression does not require the time consuming process required for the development of stable cell lines.

[0114] CHO cells are capable of successfully expressing a wide variety of proteins from a broad range of cell types. Thus, while the glycosylation pattern on a recombinant protein produced in a mammalian cell expression system may not be identical to the natural protein, the differences in oligosaccharide side chains are often not essential for biological activity of the expressed protein.

[0115] The DHFR gene also may be used as part of a gene amplification scheme for CHO cells. Another gene amplification scheme relies on the temperature sensitivity (ts) of BSC40-tsA58 cells transfected with an SV40 vector. Temperature reduction to 33 °C stabilizes the ts SV40 T antigen which leads to the excision and amplification of the integrated transfected vector DNA, thereby also amplifying the associated gene of interest.

[0116] Stable cell lines were established for CHO cells as well as BSC40-tsA58 cells (hereinafter referred to as "BSC cells"). The various cells, cell lines and DNA sequences chosen for mammalian cell expression of the Smads, serine/threonine kinase receptors and BMPs of this invention are well characterized in the art and are readily available. Other promoters, selectable markers, gene amplification methods and cells also may be used to express the Smads, serine/threonine kinase receptors and BMPs of this invention. Particular details of the

transfection, expression, and purification of recombinant proteins are well documented in the art and are understood by those having ordinary skill in the art. Further details on the various technical aspects of each of the steps used in recombinant production of foreign genes in mammalian cell expression systems can be found in a number of texts and laboratory manuals in the art. See, e.g., F. M. Ausubel et al., ed., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York (1989).

Progenitor cells

[0117] The progenitor cells that are induced to proliferate and/or differentiate in the present invention are preferably mammalian cells. Preferred progenitor cells include mammalian chondroblasts, osteoblasts and neuroblasts, all earlier developmental precursors thereof, and all cells that develop therefrom (e.g., chondroblasts, pre-chondroblasts and chondrocytes). However, any non-mammalian progenitor cells are also likely to be useful in the methods of the present invention. It is, thus, envisioned that when schemes become available for implanting xenogeneic cells into humans without causing adverse immunological reactions, non-mammalian progenitor cells will be useful for tissue regeneration and repair in humans.

[0118] In some embodiments, the progenitor cells comprise a nucleic acid encoding a Smad and optionally a nucleic acid encoding one or more serine/threonine kinase receptor and/or a nucleic acid encoding one or more BMP. In some embodiments, the progenitor cells comprise vectors comprising a nucleic acid encoding a Smad and optionally a nucleic acid encoding one or more

serine/threonine kinase receptor and/or a nucleic acid encoding one or more BMP. In some embodiments, the nucleic acid encoding Smad, serine/threonine kinase receptor and/or BMP are in one vector. In some embodiments, the nucleic acid encoding Smad, serine/threonine kinase receptor and/or BMP are in different vectors. In some embodiments, the nucleic acids are recombinant.

[0119] In some embodiments, only type I serine/threonine kinase receptors are used. In some embodiments, only type II serine/threonine kinase receptors are used. In other embodiments, both type I and type II serine/threonine kinase receptors are used. In some embodiments, more than one BMP will be administered to the desired cell or tissue. In some embodiments, two BMPS will be used. In some embodiments, three BMPs will be used. The particular choice of combination of BMPs and the relative concentrations at which they are combined may be varied systematically to optimize the tissue type induced at a selected treatment site using the procedures described herein. The relative concentrations of the BMPs either alone or in combination that will optimally induce tissue formation when administered to a mammal may be determined empirically by the skilled practitioner using the procedures described herein.

Gene Therapy

[0120] The Smad, serine/threonine kinase receptor and/or BMP proteins can be produced *in vivo* in a mammal, e.g., a human patient, using a gene therapy approach for inducing tissue formation, repairing a tissue defect or regenerating tissue at a target locus. This involves

administration of a suitable Smad, serine/threonine kinase receptor and/or BMP protein-encoding nucleic acid operably linked to suitable expression control sequences. Preferably, these sequences are incorporated into a viral vector. Suitable viral vectors for such gene therapy include adenoviral vectors, lentiviral vectors, baculoviral vectors, Epstein Barr viral vectors, papovaviral vectors, vaccinia viral vectors, herpes simplex viral vectors, and adeno associated virus (AAV) vectors. The viral vector can be a replication-defective viral vector. A preferred adenoviral vector has a deletion in its E1 gene or E3 gene. When an adenoviral vector is used, preferably the mammal is not exposed to a nucleic acid encoding a selectable marker gene.

Formulations

[0121] The proteins (e.g., BMPs) of the present invention can be formulated as part of a pharmaceutical composition. The pharmaceutical compositions provided by this invention comprise at least one protein. The compositions of this invention will be administered at an effective dose to induce the particular type of tissue at the treatment site selected according to the particular clinical condition addressed. Determination of a preferred pharmaceutical formulation and a therapeutically effective dose regimen for a given application is well within the skill of the art taking into consideration, for example, the mode of administration, the condition and weight of the patient, the extent of the desired treatment and the tolerance of the patient for the treatment.

[0122] Administration of the proteins of this invention, may be accomplished using any of the conventional modes of administration.

[0123] The pharmaceutical compositions comprising a protein of this invention may be in a variety of forms. These include, for example, solid, semi-solid and liquid dosage forms such as tablets, pills, powders, liquid solutions or suspensions, suppositories, and injectable and infusible solutions. The preferred form depends on the intended mode of administration and therapeutic application and may be selected by one skilled in the art. Modes of administration may include oral, parenteral, subcutaneous, intravenous, intralesional or topical administration. In most cases, the pharmaceutical compositions of this invention will be administered in the vicinity of the treatment site in need of tissue regeneration or repair.

[0124] The pharmaceutical compositions comprising a protein of this invention may, for example, be placed into sterile, isotonic formulations with or without cofactors which stimulate uptake or stability. The formulation is preferably liquid, or may be lyophilized powder. For example, the BMP of this invention may be diluted with a formulation buffer comprising 5.0 mg/ml citric acid monohydrate, 2.7 mg/ml trisodium citrate, 41 mg/ml mannitol, 1 mg/ml glycine and 1 mg/ml polysorbate 20. This solution can be lyophilized, stored under refrigeration and reconstituted prior to administration with sterile Water-For-Injection (USP).

[0125] The compositions also will preferably include conventional pharmaceutically acceptable carriers well known in the art (see for example Remington's Pharmaceutical Sciences, 16th Edition, 1980, Mac

Publishing Company). Such pharmaceutically acceptable carriers may include other medicinal agents, carriers, genetic carriers, adjuvants, excipients, etc., such as human serum albumin or plasma preparations. The compositions are preferably in the form of a unit dose and will usually be administered as a dose regimen that depends on the particular tissue treatment.

[0126] The pharmaceutical compositions of this invention may also be administered in conjunction with a morphogenic device using, for example, microspheres, liposomes, other microparticulate delivery systems or sustained release formulations placed in, near, or otherwise in communication with affected tissues or the bloodstream bathing those tissues.

[0127] Liposomes containing a protein of this invention can be prepared by well-known methods (See, e.g. DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. U.S.A., 82, pp. 3688-92 (1985); Hwang et al., Proc. Natl. Acad. Sci. U.S.A., 77, pp. 4030-34 (1980); U.S. Patent Nos. 4,485,045 and 4,544,545). Ordinarily the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol.% cholesterol. The proportion of cholesterol is selected to control the optimal rate of BMP release.

[0128] The proteins of this invention may also be attached to liposomes containing other biologically active molecules such as immunosuppressive agents, cytokines, etc., to modulate the rate and characteristics of tissue induction. Attachment of BMPs to liposomes may be accomplished by any known cross-linking agent such as heterobifunctional cross-linking agents that have been widely used to couple toxins or

chemotherapeutic agents to antibodies for targeted delivery. Conjugation to liposomes can also be accomplished using the carbohydrate-directed cross-linking reagent 4-(4-maleimidophenyl) butyric acid hydrazide (MPBH) (Duzgunes et al., J. Cell. Biochem. Abst. Suppl. 16E 77 (1992)).

[0129] The following are examples which illustrate the methods of this invention. These examples should not be construed as limiting: the examples are included for purposes of illustration and the present invention is limited only by the claims.

Example 1 - Effect of CDMP-1 and OP-1 on Smad5 levels

[0130] Total C2C12 cell lysates treated with solvent control (lane 1), CDMP-1 (200 ng/ml), or OP-1 (100 ng/ml) were collected after 5 days. Proteins in the lysates were resolved by SDS-PAGE, transferred to nylon membranes, and analyzed using the anti-human Smad5 antibody (Cell Signaling Technology, Inc., Beverly, MA), as the primary antibody and anti-rabbit IgG HRP-conjugated antibody was used as the secondary antibody. Immunoreactive bands were detected using the SuperSignal chemiluminescent detection system (Pierce, Rockford, IL), according to the manufacturer's instructions. Representative results of two independent experiments are shown. Western blot analyses showed that the Smad5 protein level in control C2C12 cells was hardly detectable by Western blot analysis (Figure 1, lane 1). CDMP-1 alone stimulated Smad5 protein expression in C2C12 cells (Figure 1, lane 2). OP-1 alone stimulated Smad5 protein expression to a higher level (Figure 1, lane 3). The combination of CDMP-1 and OP-1 significantly stimulated Smad5 protein expression even

more (Figure 1, lane 4). The current observation that the protein level of Smad5 was dramatically increased in cells treated with the combination of OP-1 and CDMP-1 would support the supposition that the effect of the combination of OP-1 and CDMP-1 is likely directed at the Smad signaling pathway.

Example 2 - Co-transfection with the OP-1 Gene and the CDMP-1 Gene

[0131] Target cells will be transfected with plasmids or viral carriers containing the OP-1 gene and the CDMP-1 gene under an appropriate promoter. In addition, these cells will be transfected with plasmids or viral carriers containing the Smad5 gene. For example, plasmid pW24 (10.35 kb) that contains the coding sequence for OP-1 under the control of the CMV promoter may be used. Similarly, plasmids containing the CMDP-1 gene (pCDMP-1) or the Smad5 gene (pSmad5) will be constructed by replacing the OP-1 gene in the pW24 plasmid. Alternatively, these genes may be placed under the control of promoters that can be induced thus allowing control of expression at will. Additional vectors, such as adenoviral vectors may be constructed and tested in order to optimize and maximize expression.

[0132] Target cells that have been shown capable to become osteoblastic upon treatment with OP-1 include osteoblastic cells such as FRC cells, and pluripotent mesenchymal cells such as C2C12 and C3H10T1/2 cell lines. Cells of other orthopaedic utility may also be used. These include cells derived from cartilage, tendon, ligament, and meniscus.

[0133] Transfection studies will be carried out using the calcium phosphate-DNA coprecipitation method or

FuGene6 (Roche Diagnostics). Briefly, cells will be transfected with the specified plasmid DNAs (1 μg/ml) in serum-free medium for 4-6 h. The medium will be replaced with complete αMEM containing 10% FBS.

Example 3 - Effect of OP-1, CDMP-1 and/or Smad5 on Alkaline Phosphatase Activity

[0134] The effect of co-transfection of FRC cells with pW24 and pCDMP-1 and/or pSmad5 will be examined. For example, FRC cells will be transfected with pW24, pCDMP-1 and/or pSmad5. After 48 h, total AP activity will be measured. It is anticipated that the AP activity in cells co-transfected with pW24, pCDMP-1 and/or pSmad5 will increase as a function of pCDMP-1 and/or pSmad5 concentration. The increase should be beyond that in cells transfected with pW24 alone. The AP activity in cells transfected with pCDMP-1 alone will also show a lesser but significant increase compared to the non-transfected control or cells transfected with the empty plasmids (vectors without the OP-1, CDMP-1 and/or Smad5 genes). Co-transfection with pW24 and the empty plasmid is not expected to result in an increase in AP activity beyond that by pW24 alone. Protein levels of OP-1, CDMP-1, and Smad5 under these conditions will be measured by Western blot analysis. Control experiments will include co-transfection of FRC cells with the empty plasmids, i.e. vector without the OP-1, CDMP-1 or Smad5 genes. It is expected that the AP activity will not be elevated beyond the non-transfected controls. The expression of additional osteoblastic cell markers, such as osteocalcin (OC) and bone sialoprotein (BSP) will also be monitored using either Northern blot analysis or real-time PCR. Transfection

conditions for other cell types will be optimized accordingly using art recognized methods.

Example 4 - Effect of OP-1, CDMP-1 and/or Smad5 on Mineralized Bone Nodule Formation

[0135] The effect of co-transfection of FRC cells with pW24 and pCDMP-1 or pSmad5 on mineralized bone nodule formation will be examined. Confluent FRC cells will be co-transfected with pW24 and pCDM-1 and/or pSmad5 using FuGene6 using the optimal ratio of pW24 to pCDMP-1 or pSmad5 as determined by the experiments described above. Cells will be cultured in complete αMEM containing 5% FBS, ascorbic acid, and β-glycerol phosphate in the presence of 250 μg/ml of Neomycin. Media will be changed every 3 days. Progress of nodule formation will be monitored periodically and the images will be captured using an Olympus CK2 inverted microscope equipped with a CCD camera. At the termination of the experiments, cultures will be stained by Alizarin Red-S staining to assess the extent of formation of mineralized bone nodules.

Example 5 - In vivo Expression of OP-1, CDMP-1 and Smad5

[0136] *In vivo* studies will be conducted using two experimental approaches: (i) Direct injection of OP-1 expressing vectors together with CDMP-1 and/or Smad5-expressing vectors into muscles of mice, and (ii) injection of transfected cells into muscles.

[0137] For direct injection experiments, nude mice will be injected with vectors expressing OP-1, CDMP-1 and/or Smad5 (pW24, pCDMP-1 and/or pSmad5, respectively) with a 27-gauge needle subcutaneously into a male homozygous nude mouse. Standard aseptic techniques will be used in all manipulations. To determine *in vivo*

osteogenic dose response of the vectors, eight mice will be used. Each mouse will be injected with 0.1 - 10 mg/ml vectors in 100 μl each. Body weight and growth at the site of injection will be followed daily *via* in-life measurement of the mass. The cross-sectional area of the mass will be measured with a vernier caliper. The size of the mass will be calculated using the formula: length/2 x width/2 x Π. The mass and the body weight will be plotted as a function of time following injection. The animals will be monitored for 49 days. At necropsy, the mass at the site of injection will be collected, fixed, stained with hematoxylin and eosin, and subjected to histological analysis. Controls will include mice injected with individual pW24, pCDMP-1 and/or pSmad5 alone. It is anticipated that the bone mass in mice injected with the combination of the pW24 and pCDMP-1, or pW24 plus pSmad5 will be greater than that injected with individual vector alone.

[0138] For experiments using injection of cells, similar experiments as described above will be conducted except that animals will be injected with cells co-transfected with vectors carrying the OP-1, CDMP-1, and/or Smad5 genes. Accordingly, cells will be grown to mid-log phase and transfected with a combination of vectors expressing OP-1, CDMP-1 and/or Smad5 as described above using the optimal ratio of the two vectors. Cells will be removed from the culturing dishes by trypsin-EDTA digestion. Trypsin will be inactivated by serum (10%) and removed by repeated washings with HBSS. Cells will be suspended in a minimal volume of HBSS and injected with a 27-gauge needle subcutaneously into the flank of a male homozygous nude mouse. Standard aseptic techniques will

be used in all manipulations. Eight nude mice will be injected with 10^6 cells in 100 μl each. Outcome measurements as described above will be conducted. It is anticipated that the bone mass in mice injected with cells transfected with the combination of the pW24 and pCDMP-1, or pW24 plus pSmad5 will be greater than that injected with cells transfected with individual vector alone.

Example 6 - Gene Therapy In Patients Using Transfected Cells

[0139] For cell therapeutics with tranfected genes, appropriate cells will be transfected *in vitro* with DNA vectors carrying the OP-1 gene, the CDMP-1 gene, and/or the Smad5 gene and optionally a serine/threonine kinase receptor. Appropriate cells include osteoblasts or osteoblastic cell progenitors for the repair of bone defects. For repair of cartilage regeneration, cells of chondrocyte origin or chondrogenitor cells will be appropriate. Similarly, for the regeneration of tendons or ligaments, the appropriate cells include progenitor cells of tendon or ligament origin. The transfected cells will be cultured to allow expression of the transfected gene(s). The cells will then be injected or implanted into a defect site in a patient. The defect site may be in bone, cartilage, tendon, ligament or neural tissue. The number of cells injected or implanted into the defect will depend on the size of the defect. Exemplary DNA vectors will be pW24, pCDMP-1, or pSmad5 as described previously.

Example 7 - Gene Therapy In Patients Using Transfected Cells

[0140] For directed gene therapy, a combination of vectors as described above carrying the OP-1 gene, the

cDMP-1 gene, the Smad5 gene and/or a serine/threonine kinase will be injected into the defect site in a patient. The genes encoding each of the proteins may be placed in the same vector or in separate vectors.

Example 8 - Monitoring Effects of Gene Therapy in Patients

[0141] The repair site will be monitored radiographically every two weeks for a minimum of two years. It is anticipated that the defect site which receives the combination of OP-1 + CDMP-1 + Smad or OP-1 + Smad5 (delivered by either of the methods described in Examples 6 and 7) will exhibit a faster rate of repair than that receives OP-1 alone.

[0142] In case of cartilage and tendon/ligament repair, delivery of pCDMP-1 alone to a defect site will induce tissue formation. However, delivery of OP-1 + CDMP-1 to a defect site should exhibit a faster rate of repair than delivery of CDMP-1 or Smad5 alone. In all cases, the treated sites should exhibit regeneration whereas those treated with vehicle should not.

We Claim:

1. A method of inducing the expression of a Smad in a cell or tissue comprising contacting the cell or tissue capable of expressing the Smad with a bone morphogenic protein.

2. The method according to claim 1, wherein the cell or tissue is contacted with two bone morphogenic proteins.

3. The method according to claim 1 or 2, wherein each bone morphogenic protein is independently selected from the group consisting of OP-1 (BMP-7), OP-2, OP-3, COP-1, COP-3, COP-4, COP-5, COP-7, COP-16, BMP-2, BMP-3, BMP-3b, BMP-4, BMP-5, BMP-6, BMP-9, BMP-10, BMP-11, CDMP-3 (BMP-12), CDMP-2 (BMP-13), CDMP-1 (BMP-14), BMP-15, BMP-16, BMP-17, BMP-18, GDF-1, GDF-2, GDF-3, GDF-5, GDF-6, GDF-7, GDF-8, GDF-9, GDF-10, GDF-11, GDF-12, MP121, dorsalin-1, DPP, Vg-1, Vgr-1, 60A protein, NODAL, UNIVIN, SCREW, ADMP, and NEURAL.

4. The method according to claim 1, wherein the bone morphogenic protein is OP-1.

5. The method according to claim 1, wherein the bone morphogenic protein is CDMP-1 or GDF-5.

6. The method according to claim 2, wherein the first bone morphogenic protein is OP-1 and the second bone morphogenic protein is selected from the group consisting of CDMP-1, and GDF-5.

7. The method according to claim 1 or 2, wherein the Smad is selected from the group consisting of Smad1, Smad2, Smad3, Smad5 and Smad8.

8. The method according to claim 7, wherein the Smad is Smad5.

9. The method according to claim 1, wherein the Smad is a recombinant Smad.

10. The method according to claim 1 or 2, wherein the cell or tissue is further capable of expressing a serine/threonine kinase receptor.

11. The method according to claim 10, wherein the serine/threonine kinase receptor is a type I receptor.

12. The method according to claim 10, wherein the serine/threonine kinase receptor is a type II receptor.

13. The method according to claim 10, wherein the cell or tissue is further capable of expressing both type I and type II serine/threonine kinase receptor.

14. The method according to claim 11 or 13, wherein the type I receptor is selected from the group consisting of ALK-1, ALK-2, ALK-3, ALK-4, ALK-5, ALK-6, and ALK-7.

15. The method according to claim 10 wherein the serine/threonine kinase receptor is a recombinant serine/threonine kinase receptor.

16. The method according to claim 1 or 2 wherein the tissue is selected from the group consisting of bone, cartilage, tendon, ligament and neural tissue.

17. The method according to claim 1 or 2 wherein the cell is a progenitor cell.

18. The method according to claim 17, wherein the progenitor cell is selected from the group consisting of an osteoprogenitor cell, a cartilage progenitor cell, a ligament progenitor cell, a tendon progenitor cell, and a neural progenitor cell.

19. A method of inducing tissue formation at a target locus in a mammal comprising the step of administering to the target locus a nucleic acid encoding a Smad.

20. A method of inducing tissue formation at a target locus in a mammal comprising the step of administering to the target locus a vector comprising a nucleic acid encoding a Smad operably linked to an expression control sequence.

21. A method of inducing tissue formation at a target locus in a mammal comprising the step of administering to the target locus a cell comprising a vector comprising a nucleic acid encoding a Smad operably linked to an expression control sequence.

22. A method of repairing a tissue defect or regenerating tissue at a target locus in a mammal comprising the step of administering to the target locus a nucleic acid encoding a Smad.

23. A method of repairing a tissue defect or regenerating tissue at a target locus in a mammal comprising the step of administering to the target locus a vector comprising a nucleic acid encoding a Smad operably linked to an expression control sequence.

24. A method of repairing a tissue defect or regenerating tissue at a target locus in a mammal comprising the step of administering to the target locus a cell comprising a vector comprising a nucleic acid encoding a Smad operably linked to an expression control sequence.

25. The method according to any one of claims 20, 21, 23 or 24, wherein the expression control sequence comprises a constitutive promoter.

26. The method according to claim 20, 21, 23 or 24, wherein the expression control sequence comprises an inducible promoter.

27. The method according to any one of claims 19-24, wherein the Smad is selected from the group consisting of Smad1, Smad2, Smad3, Smad5 and Smad8.

28. The method according to claim 27, wherein the Smad is Smad5.

29. The method according to any one of claims 19-28, wherein the Smad is recombinant Smad.

30. The method according to any one of claims 19-24 further comprising the step of

administering to the target locus a nucleic acid encoding a serine/threonine kinase receptor.

31. The method according to any one of claims 19-24 further comprising the step of administering to the target locus a vector comprising a nucleic acid encoding a serine/threonine kinase receptor operably linked to an expression control sequence.

32. The method according to any one of claims 19-24 further comprising the step of administering to the target locus a cell comprising a vector comprising a nucleic acid encoding a serine/threonine kinase receptor operably linked to an expression control sequence.

33. The method according to claim 31 or 32, wherein the expression control sequence operably linked to the serine/threonine kinase receptor comprises a constitutive promoter.

34. The method according to claim 31 or 32, wherein the expression control sequence operably linked to the serine/threonine kinase receptor comprises an inducible promoter.

35. The method according to any one of claims 33-32, wherein the serine/threonine kinase receptor is a type I receptor.

36. The method according to any one of claims 33-32, wherein the serine/threonine kinase receptor is a type II receptor.

37. The method according to any one of claims 33-32, wherein both type I and type II serine/threonine kinase receptors are administered.

38. The method according to any one of claim 35 or 37, wherein the type I receptor is selected from the group consisting of ALK-1, ALK-2, ALK-3, ALK-4, ALK-5, ALK-6, and ALK-7.

39. The method according to claim 38, wherein the serine/threonine kinase receptor is selected from the group consisting of ALK-2, ALK-3, and ALK-6.

40. The method according to claim 30-32, wherein the serine/threonine kinase receptor is a recombinant ALK.

41. The method according to any one of claims 19-24, further comprising the step of administering to the target locus a bone morphogenic protein.

42. The method according to any one of claims 19-24, further comprising the step of administering to the target locus a nucleic acid encoding a bone morphogenic protein.

43. The method according to any one of claims 19-24, further comprising the step of administering to the target locus a vector comprising a nucleic acid encoding a bone morphogenic protein operably linked to an expression control sequence.

44. The method according to any one of claims 19-24, further comprising the step of administering to the target locus a cell comprising a vector comprising a nucleic acid encoding a bone morphogenic protein operably linked to an expression control sequence.

45. The method according to any one of claims 41-44, wherein the bone morphogenic protein is selected from the groups consisting of OP-1 (BMP-7), OP-2, OP-3, COP-1, COP-3, COP-4, COP-5, COP-7, COP-16, BMP-2, BMP-3, BMP-3b, BMP-4, BMP-5, BMP-6, BMP-9, BMP-10, BMP-11, CDMP-3 (BMP-12), CDMP-2 (BMP-13), CDMP-1 (BMP-14), BMP-15, BMP-16, BMP-17, BMP-18, GDF-1, GDF-2, GDF-3, GDF-5, GDF-6, GDF-7, GDF-8, GDF-9, GDF-10, GDF-11, GDF-12, MP121, dorsalin-1, DPP, Vg-1, Vgr-1, 60A protein, NODAL, UNIVIN, SCREW, ADMP, and NEURAL.

46. The method according to claim 45, wherein the bone morphogenic protein is OP-1.

47. The method according to any one of claims 19-24, wherein the tissue is selected from the group consisting of bone, cartilage, tendon, ligament and neural tissue.

48. The method according to claim 21 or 24, wherein the cell is a progenitor cell.

49. The method according to claim 48, wherein the progenitor cell is selected from the group consisting of an osteoprogenitor cell, a cartilage progenitor cell, a ligament progenitor cell, a tendon progenitor cell, and a neural progenitor cell.

FIGURE 1

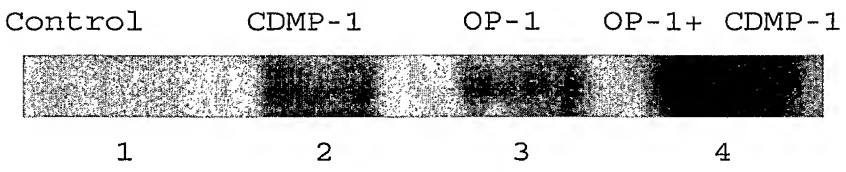
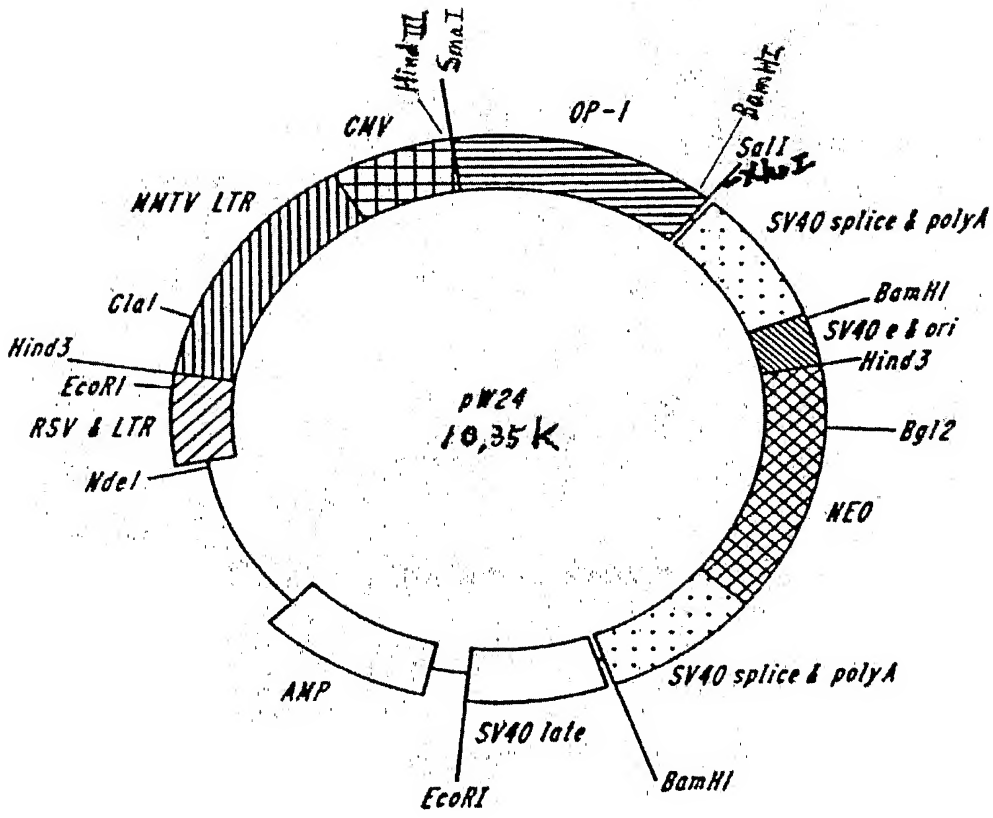


Figure 2

Plasmid map of pW24 containing OP-1 gene



SEQUENCE LISTING

<110> Stryker Corporation

<120> THERAPEUTIC METHODS USING SMADs

<130> STK-15 PCT

<140> Not yet assigned
<141> 2005-02-04

<160> 27

<170> PatentIn version 3.2

<210> 1
<211> 431
<212> PRT
<213> Homo sapiens

<400> 1

Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala
1 5 10 15

Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser
20 25 30

Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser
35 40 45

Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu
50 55 60

Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro
65 70 75 80

Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly
85 90 95

Gly Pro Gly Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser
100 105 110

Thr Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr
115 120 125

Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys
130 135 140

Glu Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu
145 150 155 160

```
Ser Lys Ile Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile
                165                 170                 175

Tyr Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Glu Arg Ile
            180                 185                 190

Ser Val Tyr Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu
        195                 200                 205

Phe Leu Leu Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu
    210                 215                 220

Val Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg
225                 230                 235                 240

His Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser
                245                 250                 255

Ile Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn
            260                 265                 270

Lys Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe
        275                 280                 285

Arg Ser Ile Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser
    290                 295                 300

Lys Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu
305                 310                 315                 320

Asn Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr
                325                 330                 335

Val Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu
            340                 345                 350

Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn
        355                 360                 365

Ser Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His
    370                 375                 380

Phe Ile Asn Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln
385                 390                 395                 400
```

Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile
405 410 415

Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
420 425 430

<210> 2
<211> 96
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic amino acid sequence COP-5

<400> 2

Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asp Asp Trp Ile Val Ala
1 5 10 15

Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His Gly Glu Cys Pro Phe Pro
20 25 30

Leu Ala Asp His Phe Asn Ser Thr Asn His Ala Val Val Gln Thr Leu
35 40 45

Val Asn Ser Val Asn Ser Lys Ile Pro Lys Ala Cys Cys Val Pro Thr
50 55 60

Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys Val
65 70 75 80

Val Leu Lys Tyr Asn Gln Glu Met Val Val Glu Gly Cys Gly Cys Arg
85 90 95

<210> 3
<211> 96
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic amino acid sequence COP-7

<400> 3

Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val Ala
1 5 10 15

Pro Pro Gly Tyr His Ala Phe Tyr Cys His Gly Glu Cys Pro Phe Pro
20 25 30

Leu Ala Asp His Leu Asn Ser Thr Asn His Ala Val Val Gln Thr Leu
35 40 45

Val Asn Ser Val Asn Ser Lys Ile Pro Lys Ala Cys Cys Val Pro Thr
50 55 60

Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys Val
65 70 75 80

Val Leu Lys Tyr Asn Gln Glu Met Val Val Glu Gly Cys Gly Cys Arg
85 90 95

<210> 4
<211> 102
<212> PRT
<213> Artificial Sequence

<220>
<223> OP-X

<220>
<221> MOD_RES
<222> (1)..(102)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (2)..(3)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (9)..(9)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (11)..(11)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (16)..(16)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (19)..(19)
<223> Xaa is independently selected from a group of one or more

specified amino acids as defined in the specification
<220>
<221> misc_feature
<222> (23)..(23)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification
<220>
<221> misc_feature
<222> (26)..(26)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification
<220>
<221> misc_feature
<222> (35)..(35)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (39)..(39)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (41)..(41)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (50)..(50)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (52)..(52)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (56)..(58)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (60)..(61)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (65)..(65)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>

<221> misc_feature
<222> (71)..(71)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (73)..(73)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (75)..(75)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (80)..(80)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (82)..(82)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (84)..(84)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (87)..(87)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (89)..(89)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (91)..(91)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (97)..(97)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<400> 4

Cys Xaa Xaa His Glu Leu Tyr Val Xaa Phe Xaa Asp Leu Gly Trp Xaa
1 5 10 15

Asp Trp Xaa Ile Ala Pro Xaa Gly Tyr Xaa Ala Tyr Tyr Cys Glu Gly
20 25 30

Glu Cys Xaa Phe Pro Leu Xaa Ser Xaa Met Asn Ala Thr Asn His Ala
35 40 45

Ile Xaa Gln Xaa Leu Val His Xaa Xaa Xaa Pro Xaa Xaa Val Pro Lys
50 55 60

Xaa Cys Cys Ala Pro Thr Xaa Leu Xaa Ala Xaa Ser Val Leu Tyr Xaa
65 70 75 80

Asp Xaa Ser Xaa Asn Val Xaa Leu Xaa Lys Xaa Arg Asn Met Val Val
85 90 95

Xaa Ala Cys Gly Cys His
100

<210> 5
<211> 97
<212> PRT
<213> Artificial Sequence

<220>
<223> Generic Sequence 7

<220>
<221> MOD_RES
<222> (1)..(97)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (2)..(4)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (6)..(8)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (11)..(16)
<223> Xaa is independently selected from a group of one or more

specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (18)..(21)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (23)..(23)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (26)..(26)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (28)..(28)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (30)..(31)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (33)..(40)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (44)..(60)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (63)..(63)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (65)..(72)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (74)..(80)
<223> Xaa is independently selected from a group of one or more

specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (82)..(82)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (84)..(88)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (90)..(90)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (92)..(93)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (95)..(95)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (97)..(97)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<400> 5

Leu Xaa Xaa Xaa Phe Xaa Xaa Xaa Gly Trp Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Pro Xaa Xaa Xaa Xaa Ala Xaa Tyr Cys Xaa Gly Xaa Cys Xaa Xaa Pro
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asn His Ala Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Cys Xaa Pro
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65 70 75 80

Val Xaa Leu Xaa Xaa Xaa Xaa Xaa Met Xaa Val Xaa Xaa Cys Xaa Cys
85 90 95

Xaa

<210> 6
<211> 102
<212> PRT
<213> Artificial Sequence

<220>
<223> Generic Sequence 8

<220>
<221> MOD_RES
<222> (1)..(102)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (2)..(5)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (7)..(9)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (11)..(13)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (16)..(21)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (23)..(26)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (28)..(28)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>

<221> misc_feature
<222> (31)..(31)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (33)..(33)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (35)..(36)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (38)..(45)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (49)..(65)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (68)..(68)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (70)..(77)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (79)..(85)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (87)..(87)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (89)..(93)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>

<221> misc_feature
<222> (95)..(95)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (97)..(98)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (100)..(100)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (102)..(102)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<400> 6

Cys Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Phe Xaa Xaa Xaa Gly Trp Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Pro Xaa Xaa Xaa Xaa Ala Xaa Tyr Cys Xaa Gly
20 25 30

Xaa Cys Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asn His Ala
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
50 55 60

Xaa Cys Cys Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Val Xaa Leu Xaa Xaa Xaa Xaa Xaa Met Xaa Val
85 90 95

Xaa Xaa Cys Xaa Cys Xaa
100

<210> 7
<211> 97
<212> PRT
<213> Artificial Sequence

<220>
<223> Generic Sequence 9

<220>
<221> MOD_RES
<222> (1)..(97)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (1)..(16)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (18)..(24)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (26)..(26)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (28)..(28)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (30)..(61)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (63)..(63)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (65)..(72)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (74)..(93)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (95)..(95)
<223> Xaa is independently selected from a group of one or more

specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (97)..(97)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<400> 7

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Gly Xaa Cys Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Pro
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Cys
85 90 95

Xaa

<210> 8
<211> 102
<212> PRT
<213> Artificial Sequence

<220>
<223> Generic Sequence 10

<220>
<221> MOD_RES
<222> (1)..(102)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (2)..(21)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>

<221> misc_feature
<222> (23)..(29)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (31)..(31)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (33)..(33)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (35)..(66)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (68)..(68)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (70)..(77)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (79)..(98)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (100)..(100)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (102)..(102)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<400> 8

Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Gly

20 25 30

Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
50 55 60

Xaa Xaa Cys Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
85 90 95

Xaa Xaa Cys Xaa Cys Xaa
100

<210> 9
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Consensus Sequence

<220>
<221> MOD_RES
<222> (1)..(5)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<221> misc_feature
<222> (2)..(5)
<223> Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<400> 9

Cys Xaa Xaa Xaa Xaa
1 5

<210> 10
<211> 1822
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (49)..(1341)

<400> 10

```
ggtgcgggcc cggagcccgg agcccgggta gcgcgtagag ccggcgcg atg cac gtg      57
                                                        Met His Val
                                                        1

cgc tca ctg cga gct gcg gcg ccg cac agc ttc gtg gcg ctc tgg gca      105
Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala Leu Trp Ala
    5                   10                  15

ccc ctg ttc ctg ctg cgc tcc gcc ctg gcc gac ttc agc ctg gac aac      153
Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser Leu Asp Asn
20                  25                  30                  35

gag gtg cac tcg agc ttc atc cac cgg cgc ctc cgc agc cag gag cgg      201
Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser Gln Glu Arg
                40                  45                  50

cgg gag atg cag cgc gag atc ctc tcc att ttg ggc ttg ccc cac cgc      249
Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg
            55                  60                  65

ccg cgc ccg cac ctc cag ggc aag cac aac tcg gca ccc atg ttc atg      297
Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro Met Phe Met
        70                  75                  80

ctg gac ctg tac aac gcc atg gcg gtg gag gag ggc ggc ggg ccc ggc      345
Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly Gly Pro Gly
    85                  90                  95

ggc cag ggc ttc tcc tac ccc tac aag gcc gtc ttc agt acc cag ggc      393
Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr Gln Gly
100                 105                 110                 115

ccc cct ctg gcc agc ctg caa gat agc cat ttc ctc acc gac gcc gac      441
Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp Ala Asp
                120                 125                 130

atg gtc atg agc ttc gtc aac ctc gtg gaa cat gac aag gaa ttc ttc      489
Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu Phe Phe
            135                 140                 145

cac cca cgc tac cac cat cga gag ttc cgg ttt gat ctt tcc aag atc      537
His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser Lys Ile
        150                 155                 160

cca gaa ggg gaa gct gtc acg gca gcc gaa ttc cgg atc tac aag gac      585
Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp
    165                 170                 175

tac atc cgg gaa cgc ttc gac aat gag acg ttc cgg atc agc gtt tat      633
Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile Ser Val Tyr
180                 185                 190                 195

cag gtg ctc cag gag cac ttg ggc agg gaa tcg gat ctc ttc ctg ctc      681
Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu Phe Leu Leu
                200                 205                 210

gac agc cgt acc ctc tgg gcc tcg gag gag ggc tgg ctg gtg ttt gac      729
Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu Val Phe Asp
            215                 220                 225
```

```
atc aca gcc acc agc aac cac tgg gtg gtc aat ccg cgg cac aac ctg      777
Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His Asn Leu
        230                 235                 240

ggc ctg cag ctc tcg gtg gag acg ctg gat ggg cag agc atc aac ccc      825
Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile Asn Pro
    245                 250                 255

aag ttg gcg ggc ctg att ggg cgg cac ggg ccc cag aac aag cag ccc      873
Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys Gln Pro
260                 265                 270                 275

ttc atg gtg gct ttc ttc aag gcc acg gag gtc cac ttc cgc agc atc      921
Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe Arg Ser Ile
                280                 285                 290

cgg tcc acg ggg agc aaa cag cgc agc cag aac cgc tcc aag acg ccc      969
Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro
            295                 300                 305

aag aac cag gaa gcc ctg cgg atg gcc aac gtg gca gag aac agc agc     1017
Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu Asn Ser Ser
        310                 315                 320

agc gac cag agg cag gcc tgt aag aag cac gag ctg tat gtc agc ttc     1065
Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe
    325                 330                 335

cga gac ctg ggc tgg cag gac tgg atc atc gcg cct gaa ggc tac gcc     1113
Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala
340                 345                 350                 355

gcc tac tac tgt gag ggg gag tgt gcc ttc cct ctg aac tcc tac atg     1161
Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met
                360                 365                 370

aac gcc acc aac cac gcc atc gtg cag acg ctg gtc cac ttc atc aac     1209
Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn
            375                 380                 385

ccg gaa acg gtg ccc aag ccc tgc tgt gcg ccc acg cag ctc aat gcc     1257
Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala
        390                 395                 400

atc tcc gtc ctc tac ttc gat gac agc tcc aac gtc atc ctg aag aaa     1305
Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys
    405                 410                 415

tac aga aac atg gtg gtc cgg gcc tgt ggc tgc cac tagctcctcc          1351
Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
420                 425                 430

gagaattcag acccttgggg gccaagtttt tctggatcct ccattgctcg ccttggccag   1411

gaaccagcag accaactgcc ttttgtgaga ccttcccctc cctatcccca actttaaagg   1471

tgtgagagta ttaggaaaca tgagcagcat atggcttttg atcagttttt cagtggcagc   1531

atccaatgaa caagatccta caagctgtgc aggcaaaacc tagcaggaaa aaaaaacaac   1591
```

gcataaagaa aaatggccgg gccaggtcat tggctgggaa gtctcagcca tgcacggact 1651

cgtttccaga ggtaattatg agcgcctacc agccaggcca cccagccgtg ggaggaaggg 1711

ggcgtggcaa ggggtgggca cattggtgtc tgtgcgaaag gaaaattgac ccggaagttc 1771

ctgtaataaa tgtcacaata aaacgaatga atgaaaaaaa aaaaaaaaaa a 1822

<210> 11
<211> 431
<212> PRT
<213> Homo sapiens

<400> 11

Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala
1 5 10 15

Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser
20 25 30

Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser
35 40 45

Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu
50 55 60

Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro
65 70 75 80

Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly
85 90 95

Gly Pro Gly Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser
100 105 110

Thr Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr
115 120 125

Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys
130 135 140

Glu Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu
145 150 155 160

Ser Lys Ile Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile
165 170 175

Tyr Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile
180 185 190

Ser Val Tyr Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu
195 200 205

Phe Leu Leu Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu
210 215 220

Val Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg
225 230 235 240

His Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser
245 250 255

Ile Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn
260 265 270

Lys Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe
275 280 285

Arg Ser Ile Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser
290 295 300

Lys Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu
305 310 315 320

Asn Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr
325 330 335

Val Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu
340 345 350

Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn
355 360 365

Ser Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His
370 375 380

Phe Ile Asn Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln
385 390 395 400

Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile
405 410 415

Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
420 425 430

<210> 12
<211> 3151
<212> DNA
<213> Mus musculus

<400> 12
ggaggctggg cggttccgac tgcctgacgc cagcgcgacc agatcaatcc aggctcgggg 60

agcgagcggg cgcaccaagg cgaggccggg gccgaggcgc ggggacggcg gcccggagct 120

aagcagagcg cggggacggc ggccgggagc ggatcggagc acgggacccg gcgccgggtc 180

tcgtgcgtcc ctgcggatgg gcgcgccgcc gagccggcgc taactgggat cctcgctgga 240

acaggaggga cagtattttc tacctttcca aaccgcagac caagaagcta aggagaatct 300

atgtaaatat actgaaatct ctgttggctc tgcgcccaac accccggagc tggcacctca 360

ccctgtctga ggagcgtgta gaactagacc agccgctatg aatgtgacca gcttgttttc 420

attcacaagt ccagctgtga agagactcct tgggtggaaa cagggcgatg aagaagagaa 480

atgggcagag aaagctgtgg acgctttggt gaagaaactg aagaagaaga aaggggccat 540

ggaagagctg gagaaggccc tgagctgccc tggacagccg agtaactgcg tcaccattcc 600

tcgctccctg gatggcaggt tgcaggtgtc ccaccggaag ggactacctc atgtcattta 660

ttgccgtgtg tggcgctggc ccgacctcca gagccaccat gaactgaagc ctctggaatg 720

ctgtgagttc ccatttggtt ccaagcagaa ggaggtctgc atcaacccct accactataa 780

gcgagtggag agcccggttc tcccgccggt gctggttccg aggcacagcg agtacaatcc 840

tcagcacagc cttctggctc agttccgcaa cctgggacaa aatgagcctc acatgccact 900

gaacgccacg ttcccagact ctttccagca gcccaacagc cacccgttcc cccactcccc 960

caacagcagc taccccaact ctcctggcgg cagcagcagc acctaccctc actccccaac 1020

cagctcagac ccgggcagcc cttttcagat gccagctgac acacccccac ctgcttacct 1080

gcctcctgaa gaccccatgg cccaggatgg ctctcagccc atggacacga acatgatggc 1140

gcctccactg cccgctgaaa tcagcagagg agatgttcag gcagttgctt acgaggaacc 1200

aaaacactgg tgctctattg tgtactatga gctcaacaac cgtgtgggtg aagcgttcca 1260

cgcctcgtcc accagcgtgc tggtggatgg tttcacagat ccgtccaaca ataagaaccg 1320

cttctgcctt ggcttgctct ccaacgttaa ccggaattcc actattgaaa acaccaggcg 1380

acatattggg aaaggagtcc acctttatta cgttggagga gaggtgtatg cggaatgcct 1440

cagtgacagc agcatcttcg tgcagagccg gaactgcaac taccaccacg gctttcaccc 1500

caccaccgtc tgcaagatcc ccagcgggtg cagcttgaaa atcttcaaca accaagagtt 1560

tgctcagcta ctggcgcagt ctgtgaacca cgggttcgag accgtgtatg aactcaccaa 1620

aatgtgcact attcggatga gcttcgtgaa gggttgggga gccgaatacc accggcagga 1680

tgttaccagc acccccctgct ggattgagat ccatctgcat ggccctctcc agtggctgga 1740

taaggttctg acccagatgg gctcacccca caatcctatt tcatccgtgt cttaaaagac 1800

ctgtggcttc cgtctcttgc aaactatcga gccttgcatg tacttgaagg atggacaagt 1860

cagacaggat ggagacctga cgaaggagcc acgataatac ttgacctctg tgaccaacta 1920

ttggattgag aaactgacaa gccttggttg atagcaagaa ccctttcagt ttacattgtg 1980

acattctgtt gtaaaaatca actaaaatgc tgactttcag caggactttt gtgtatagtt 2040

aaaaaaaaaa gagatggcca agccagggac aaattatcta ttaggaaaaa agaaaaaaat 2100

gattgtaatc aatccttttg tgtggggtgt tggcagaagg ttggcgctga tagtctttct 2160

gaagtgggct ttcatcaggc tcagagccca cgttgaatca tcttctcatg ggttttctta 2220

atattttaaa actacttgtt tagaaatgaa tgggtttttt gtttgttttt aaagtacagg 2280

ttaatcgtta tgacatgcat agtaatcttt ctgaaactgt atgctggctg tattactgtc 2340

agaatgatgg caggcatatg ctctttgcta aatatgtata tacagaatat ttggaggtta 2400

tgaatagtct aaatggctag tgggtttaca gagtatctga ggggcggggt cgggaagaaa 2460

acgacggctg caaatgtaga ctataccgta aagctcagct tgctgcctta aacagacaag 2520

ctggtgtctg aatttgctgt gtttcagttt ttgtagagtt ttatctgact tcttttcttc 2580

tgtcttatcc gctccacggc acagttaagc agctggttaa ttcctctaac tgtgagagca 2640

gatgagtaat tccttctgtt cgcaaatcaa ctggcttcgt gtttcagtac ccagtatatg 2700

aaaagcttga attgaatgag cagtttttat ggagtttaca gtacagacat aggctttgat 2760

ttccaaataa attgtttgcc aaacctggta actctgttca ttattcgcag gattaaagat 2820

ctctctattg gaatccattt caaaggttgt ttttttgtt tttgtttttg ttttttgttt 2880

tattttgatt tgttttttt tgtactattt ggtttctttt cttctgttaa ttttttttatt 2940

ctcctttgct cttatacagc gagtactttt attccaacac tagcagggtt tttctctact 3000

ggaaattttt aaataaaacc tgtcattatt gcttactttg attaaaaaaa aaaaaaaaaa 3060

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3120

aaaaaaaaaa aaaatttcat ctgtgtctta a 3151

<210> 13
<211> 465

<212> PRT
<213> Mus musculus

<400> 13

Met Asn Val Thr Ser Leu Phe Ser Phe Thr Ser Pro Ala Val Lys Arg
1 5 10 15

Leu Leu Gly Trp Lys Gln Gly Asp Glu Glu Glu Lys Trp Ala Glu Lys
20 25 30

Ala Val Asp Ala Leu Val Lys Lys Leu Lys Lys Lys Lys Gly Ala Met
35 40 45

Glu Glu Leu Glu Lys Ala Leu Ser Cys Pro Gly Gln Pro Ser Asn Cys
50 55 60

Val Thr Ile Pro Arg Ser Leu Asp Gly Arg Leu Gln Val Ser His Arg
65 70 75 80

Lys Gly Leu Pro His Val Ile Tyr Cys Arg Val Trp Arg Trp Pro Asp
85 90 95

Leu Gln Ser His His Glu Leu Lys Pro Leu Glu Cys Cys Glu Phe Pro
100 105 110

Phe Gly Ser Lys Gln Lys Glu Val Cys Ile Asn Pro Tyr His Tyr Lys
115 120 125

Arg Val Glu Ser Pro Val Leu Pro Pro Val Leu Val Pro Arg His Ser
130 135 140

Glu Tyr Asn Pro Gln His Ser Leu Leu Ala Gln Phe Arg Asn Leu Gly
145 150 155 160

Gln Asn Glu Pro His Met Pro Leu Asn Ala Thr Phe Pro Asp Ser Phe
165 170 175

Gln Gln Pro Asn Ser His Pro Phe Pro His Ser Pro Asn Ser Ser Tyr
180 185 190

Pro Asn Ser Pro Gly Gly Ser Ser Ser Thr Tyr Pro His Ser Pro Thr
195 200 205

Ser Ser Asp Pro Gly Ser Pro Phe Gln Met Pro Ala Asp Thr Pro Pro
210 215 220

```
Pro Ala Tyr Leu Pro Pro Glu Asp Pro Met Ala Gln Asp Gly Ser Gln
225                 230                 235                 240

Pro Met Asp Thr Asn Met Met Ala Pro Pro Leu Pro Ala Glu Ile Ser
                245                 250                 255

Arg Gly Asp Val Gln Ala Val Ala Tyr Glu Glu Pro Lys His Trp Cys
            260                 265                 270

Ser Ile Val Tyr Tyr Glu Leu Asn Asn Arg Val Gly Glu Ala Phe His
        275                 280                 285

Ala Ser Ser Thr Ser Val Leu Val Asp Gly Phe Thr Asp Pro Ser Asn
    290                 295                 300

Asn Lys Asn Arg Phe Cys Leu Gly Leu Leu Ser Asn Val Asn Arg Asn
305                 310                 315                 320

Ser Thr Ile Glu Asn Thr Arg Arg His Ile Gly Lys Gly Val His Leu
                325                 330                 335

Tyr Tyr Val Gly Gly Glu Val Tyr Ala Glu Cys Leu Ser Asp Ser Ser
            340                 345                 350

Ile Phe Val Gln Ser Arg Asn Cys Asn Tyr His His Gly Phe His Pro
        355                 360                 365

Thr Thr Val Cys Lys Ile Pro Ser Gly Cys Ser Leu Lys Ile Phe Asn
    370                 375                 380

Asn Gln Glu Phe Ala Gln Leu Leu Ala Gln Ser Val Asn His Gly Phe
385                 390                 395                 400

Glu Thr Val Tyr Glu Leu Thr Lys Met Cys Thr Ile Arg Met Ser Phe
                405                 410                 415

Val Lys Gly Trp Gly Ala Glu Tyr His Arg Gln Asp Val Thr Ser Thr
            420                 425                 430

Pro Cys Trp Ile Glu Ile His Leu His Gly Pro Leu Gln Trp Leu Asp
        435                 440                 445

Lys Val Leu Thr Gln Met Gly Ser Pro His Asn Pro Ile Ser Ser Val
    450                 455                 460
```

Ser
465

<210> 14
<211> 2449
<212> DNA
<213> Mus musculus

<400> 14
ggttaaaata actatctgag atttgttttg ctgttgttgt tgtttaagga aaattaaggt 60

agtaccatat cttaaatcat tgcaacaaga ggcagtattg ctacttataa aagtaaataa 120

tagtgtataa aattgtgttt caaccgaatc ttactggcat ctttctctct ttcttggaaa 180

cactccatga aacaatagat gcagtagatc aggatgatga ggacgggaat gggggcacta 240

ctacactact atactactac actctaggat tcgaggctgc atgcagagtt aacaacagtc 300

agctgactgt ttacctgaaa gactggcata gaataggaaa atttggtgcc aagtgcataa 360

aaataagcaa atgaaaagac attaattctg ggtagattta ccgggctttt tctgagtgtg 420

gattgttacc tttggtaaga aaatgtcgtc catcttgcca ttcactccgc cagtggtgaa 480

gagacttctg ggatggaaaa aatcagccgg tgggtctgga ggagcaggtg gtggagagca 540

gaatggacag gaagaaaagt ggtgtgaaaa agcagtgaaa agtctggtga aaaagctaaa 600

gaaaacagga cggttagatg agcttgagaa agccatcacc actcagaatt gcaatactaa 660

atgtgtcacc ataccaagca cttgctctga aatttgggga ctgagtacag caaatacggt 720

agatcagtgg gacacaacag gcctttacag cttctctgaa caaaccaggt ctcttgatgg 780

ccgtcttcag gtttcacacc ggaaagggtt gccacatgtt atatattgcc ggctctggcg 840

ctggccggac cttcacagtc atcatgagct caaggcaatc gaaaactgcg aatatgcttt 900

taatctgaaa aaagatgaag tgtgtgtaaa tccgtaccac taccagagag ttgagacccc 960

agtcttgcct ccagtcttag tgcctcggca cacggagatt ctaacagaac tgccaccct 1020

ggatgactac acccactcca ttccagaaaa cacaaatttc ccagcaggaa ttgagccaca 1080

gagtaattac atcccagaaa caccaccacc tggatatatc agtgaagatg gagaaacaag 1140

tgaccaacag ttgaaccaaa gtatggacac aggctctccg gctgaactgt ctcctactac 1200

tctctctcct gttaatcaca gcttggattt gcagccagtt acttactcgg aacctgcatt 1260

ctggtgttca atcgcatact atgaactaaa ccagagggtt ggagagacct tccatgcgtc 1320

acagccctcg ctcactgtag acggcttcac agacccatca aactcggaga ggttctgctt 1380

aggcttgctc tccaacgtta accgaaatgc cactgtagaa atgacaagaa gacatatagg 1440

aaggggagtg cgcttgtatt acataggtgg ggaagtgttt gctgagtgcc taagtgatag 1500

tgcaatcttt gtgcagagcc ccaactgtaa ccagagatac ggctggcacc ctgcaacagt 1560

gtgtaagatc ccaccaggct gtaacctgaa gatcttcaac aaccaagaat ttgctgctct 1620

tctggctcag tctgtcaacc agggttttga agccgtttat cagctaaccc gaatgtgcac 1680

cataagaatg agttttgtga agggctgggg agcagaatat cggaggcaga cagtaacaag 1740

tactccttgc tggattgaac ttcatctgaa tggccctctg cagtggctgg acaaagtatt 1800

aactcagatg ggatcccctt cagtgcgatg ctcaagcatg tcgtaaaccc atcaaagact 1860

cgctgtaaca gctcctccgt cgtagtattc atgtatgatc ccgtggactg tttgctatcc 1920

aaaaattcca gagcaaaaac agcacttgag gtctcatcag ttaaagcacc ttgtggaatc 1980

tgtttcctat atttgaatat tagatgggaa aattagtgtc tagaaatgcc ctccccagcg 2040

gggaaaaaga agacttaaag acttaatgat gtcttgttgg gcataagaca gtatcccaaa 2100

ggttattaat aacagtagta gttgtgtaca ggtaatgtgt ccagacccag tattgcagta 2160

ctatgctgtt tgtatacatt cttagtttgc ataaatgagg tgtgtgtgct gcttcttggt 2220

ctaggcaagc ctttataaaa ttacagtatc taatctgtta ttcccacttc tccgttattt 2280

ttgtgtcttt tttaatatat aatatatata tatcaagatt ttcaaattat catttagaag 2340

cagattttcc ttgtagaaac taatttttct gccttttacc aaaaataaac aaactcttgg 2400

gggaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaa 2449

<210> 15
<211> 494
<212> PRT
<213> Mus musculus

<400> 15

Met Ser Ser Ile Leu Pro Phe Thr Pro Pro Val Val Lys Arg Leu Leu
1 5 10 15

Gly Trp Lys Lys Ser Ala Gly Gly Ser Gly Gly Ala Gly Gly Gly Glu
20 25 30

Gln Asn Gly Gln Glu Glu Lys Trp Cys Glu Lys Ala Val Lys Ser Leu
35 40 45

Val Lys Lys Leu Lys Lys Thr Gly Arg Leu Asp Glu Leu Glu Lys Ala
50 55 60

Ile Thr Thr Gln Asn Cys Asn Thr Lys Cys Val Thr Ile Pro Ser Thr
65 70 75 80

```
Cys Ser Glu Ile Trp Gly Leu Ser Thr Ala Asn Thr Val Asp Gln Trp
                85                  90                  95

Asp Thr Thr Gly Leu Tyr Ser Phe Ser Glu Gln Thr Arg Ser Leu Asp
            100                 105                 110

Gly Arg Leu Gln Val Ser His Arg Lys Gly Leu Pro His Val Ile Tyr
        115                 120                 125

Cys Arg Leu Trp Arg Trp Pro Asp Leu His Ser His His Glu Leu Lys
    130                 135                 140

Ala Ile Glu Asn Cys Glu Tyr Ala Phe Asn Leu Lys Lys Asp Glu Val
145                 150                 155                 160

Cys Val Asn Pro Tyr His Tyr Gln Arg Val Glu Thr Pro Val Leu Pro
                165                 170                 175

Pro Val Leu Val Pro Arg His Thr Glu Ile Leu Thr Glu Leu Pro Pro
            180                 185                 190

Leu Asp Asp Tyr Thr His Ser Ile Pro Glu Asn Thr Asn Phe Pro Ala
        195                 200                 205

Gly Ile Glu Pro Gln Ser Asn Tyr Ile Pro Glu Thr Pro Pro Pro Gly
    210                 215                 220

Tyr Ile Ser Glu Asp Gly Glu Thr Ser Asp Gln Gln Leu Asn Gln Ser
225                 230                 235                 240

Met Asp Thr Gly Ser Pro Ala Glu Leu Ser Pro Thr Thr Leu Ser Pro
                245                 250                 255

Val Asn His Ser Leu Asp Leu Gln Pro Val Thr Tyr Ser Glu Pro Ala
            260                 265                 270

Phe Trp Cys Ser Ile Ala Tyr Tyr Glu Leu Asn Gln Arg Val Gly Glu
        275                 280                 285

Thr Phe His Ala Ser Gln Pro Ser Leu Thr Val Asp Gly Phe Thr Asp
    290                 295                 300

Pro Ser Asn Ser Glu Arg Phe Cys Leu Gly Leu Leu Ser Asn Val Asn
305                 310                 315                 320
```

Arg Asn Ala Thr Val Glu Met Thr Arg Arg His Ile Gly Arg Gly Val
325 330 335

Arg Leu Tyr Tyr Ile Gly Gly Glu Val Phe Ala Glu Cys Leu Ser Asp
340 345 350

Ser Ala Ile Phe Val Gln Ser Pro Asn Cys Asn Gln Arg Tyr Gly Trp
355 360 365

His Pro Ala Thr Val Cys Lys Ile Pro Pro Gly Cys Asn Leu Lys Ile
370 375 380

Phe Asn Asn Gln Glu Phe Ala Ala Leu Leu Ala Gln Ser Val Asn Gln
385 390 395 400

Gly Phe Glu Ala Val Tyr Gln Leu Thr Arg Met Cys Thr Ile Arg Met
405 410 415

Ser Phe Val Lys Gly Trp Gly Ala Glu Tyr Arg Arg Gln Thr Val Thr
420 425 430

Ser Thr Pro Cys Trp Ile Glu Leu His Leu Asn Gly Pro Leu Gln Trp
435 440 445

Leu Asp Lys Val Leu Thr Gln Met Gly Ser Pro Ser Val Arg Cys Ser
450 455 460

Ser Met Ser Glu Ile Thr Ala Cys Ile Ser Gln Arg Gln Val Lys Glu
465 470 475 480

Leu Asp Val Leu Ile Gln Thr Glu Gln Trp Pro Gln Tyr Pro
485 490

<210> 16
<211> 6256
<212> DNA
<213> Homo sapiens

<400> 16
gcggccgccg cctccgcccc gcgttcgggg ccttcccgac cctgcactgc tgccgtccgc 60

ccgcccggcc gctcttctct tcgccgtggg agccgctccg ggcgcagggc cgcgcgccga 120

gccccgcagg ctgcagcgcc gcggcccggc ccggcgcccc ggcaacttcg ccgagagttg 180

aggcgaagtt tgggcgaccg cggcaggccc cggccgagct cccctctgcg cccccggcgt 240

cccgtcgagc ccagccccgc cgggggcgct cctcgccgcc cgcgcgccct ccccagccat 300

gtcgtccatc ctgcctttca ctcccccgat cgtgaagcgc ctgctgggct ggaagaaggg 360

cgagcagaac gggcaggagg agaaatggtg cgagaaggcg gtcaagagcc tggtcaagaa 420

actcaagaag acggggcagc tggacgagct ggagaaggcc atcaccacgc agaacgtcaa 480

caccaagtgc atcaccatcc ccaggtccct ggatggccgg ttgcaggtgt cccatcggaa 540

ggggctccct catgtcatct actgccgcct gtggcgatgg ccagacctgc acagccacca 600

cgagctacgg gccatggagc tgtgtgagtt cgccttcaat atgaagaagg acgaggtctg 660

cgtgaatccc taccactacc agagagtaga gacaccagtt ctacctcctg tgttggtgcc 720

acgccacaca gagatcccgg ccgagttccc cccactggac gactacagcc attccatccc 780

cgaaaacact aacttccccg caggcatcga gccccagagc aatattccag agaccccacc 840

ccctggctac ctgagtgaag atggagaaac cagtgaccac cagatgaacc acagcatgga 900

cgcaggttct ccaaacctat ccccgaatcc gatgtcccca gcacataata acttggacct 960

gcagccagtt acctactgcg agccggcctt ctggtgctcc atctcctact acgagctgaa 1020

ccagcgcgtc ggggagacat tccacgcctc gcagccatcc atgactgtgg atggcttcac 1080

cgacccctcc aattcggagc gcttctgcct agggctgctc tccaatgtca acaggaatgc 1140

agcagtggag ctgacacgga gacacatcgg aagaggcgtg cggctctact acatcggagg 1200

ggaggtcttc gcagagtgcc tcagtgacag cgctattttt gtccagtctc ccaactgtaa 1260

ccagcgctat ggctggcacc cggccaccgt ctgcaagatc ccaccaggat gcaacctgaa 1320

gatcttcaac aaccaggagt tcgctgccct cctggcccag tcggtcaacc agggctttga 1380

ggctgtctac cagttgaccc gaatgtgcac catccgcatg agcttcgtca aaggctgggg 1440

agcggagtac aggagacaga ctgtgaccag tacccccctgc tggattgagc tgcacctgaa 1500

tgggcctttg cagtggcttg acaaggtcct cacccagatg ggctccccaa gcatccgctg 1560

ttccagtgtg tcttagagac atcaagtatg gtaggggagg gcaggcttgg ggaaaatggc 1620

catgcaggag gtggagaaaa ttggaactct actcaaccca ttgttgtcaa ggaagaagaa 1680

atctttctcc ctcaactgaa ggggtgcacc cacctgtttt ctgaaacaca cgagcaaacc 1740

cagaggtgga tgttatgaac agctgtgtct gccaaacaca tttacccttt ggccccactt 1800

tgaagggcaa gaaatggcgt ctgctctggt ggcttaagtg agcagaacag gtagtattac 1860

accaccggcc ccctcccccc agactctttt tttgagtgac agctttctgg gatgtcacag 1920

tccaaccaga aacacccctc tgtctaggac tgcagtgtgg agttcacctt ggaagggcgt 1980

tctaggtagg aagagcccgc agggccatgc agacctcatg cccagctctc tgacgcttgt 2040

gacagtgcct cttccagtga acattcccag cccagccccg ccccgccccg ccccaccact 2100

ccagcagacc ttgccccttg tgagctggat agacttggga tggggaggga gggagttttg 2160

tctgtctccc tcccctctca gaacatactg attgggaggt gcgtgttcag cagaacctgc 2220

acacaggaca gcgggaaaaa tcgatgagcg ccacctcttt aaaaactcac ttacgtttgt 2280

ccttttttcac tttgaaaagt tggaaggatc tgctgaggcc cagtgcatat gcaatgtata 2340

gtgtctatta tcacattaat ctcaaagaga ttcgaatgac ggtaagtgtt ctcatgaagc 2400

aggaggccct tgtcgtggga tggcatttgg tctcaggcag caccacactg ggtgcgtctc 2460

cagtcatctg taagagcttg ctccagattc tgatgcatac ggctatattg gtttatgtag 2520

tcagttgcat tcattaaatc aactttatca tatgctcttt taaatgtttg gtttatatat 2580

tttctttaaa aatcctgggc tggcacattg actgggaaac ctgagtgaga cccagcaact 2640

gcttctctcc cttctctctc ctgaggtgaa gcttttccag gttttgttga agagatacct 2700

gccagcactt ctgcaagctg aaatttacag aagcaaattc accagaaggg aaacatctca 2760

ggccaacata ggcaaatgaa aagggctatt aaaatatttt tacacctttg aaaattgcag 2820

gcttggtaca aagaggtctg tcatcttccc cctgggatat aagatgatct agctcccggt 2880

agaggatcac cggtgacaac tatagcagtt gtattgtgta acaagtactg ctcccagcag 2940

caattaggga gaaaactagt ctaaattatt tcaactggaa aaaagaaaaa agagtcctct 3000

tcttttccca gccttttgca gaacacagta gacagaactg ccaccttcaa ttggtacttt 3060

attctttgct gctgtttttg tataaaatga cctatcccac gtttttgcat gaatttatag 3120

caggaaaaat caagggattt cctatggaag tcctgcttta ttccaggtga agggaaggaa 3180

gtgtatatac ttttggcaag tcatacagct caaatgtgat gagatttctg atgttagagg 3240

gagatggaga ggcttcctga tgcctcatct gcagggtcct gtgcctctga agttctagcc 3300

atgaggtttc caggtaggac agctgctccc caagcctcct gaggacacag gaagagacgg 3360

aaggagcacc ttgacagact tgtgtgagtc ttctcgaagg agggttgact cagaacccag 3420

agacaataca aaacccctca cttcctctga gagggccaaa tgctgtgagt ctgaagtatg 3480

tgcctggtgt gaaatgatct atggcctgtt tcttacacag gaagcccccc gaacctcctg 3540

tacatgtgtt catgttccca gccagctctg agacccagga accaaatatt ccattttggc 3600

ttctgctaga gcagtcatgg ttcctctcct aaaagccatg ggcagcagtt tccgagggcc 3660

tgcatgatcc acctgctgca cgatcctatg agggcttcct gtggcacaca gccctctggg 3720

tgcttgggaa ctagcttcag gcacagcctg attctggtga tccagtgatc tatggaagtc 3780

gtgtcttact ccaggtgaag ggggaaaaaa aaagcctata ctttggcagg ttatgaactt 3840

tgaatgtgat gaaatgacac gtttggctgc atttggatgg tgtcttagaa ccctcattgc 3900

tcagacctga aggctacttc taggagcatg aagtttgagt tttgtgtttt tccaaaggat 3960

acttccttgg ccctttttct ttattgacta gaccaccaga ggaggatgtg tgggattgta 4020

ggcaaaccca cctgtggcat cactgaaaat aaatttgatc atacctaaga ggttaggaaa 4080

tggtgccatt cccaccttag agtgctacat aggtgctttg ggcgtatgta acattagtgt 4140

ccttccttga agccacaagc tagttttctt agttttaaaa tcctgttgta tgaatggcat 4200

ttgtatatta aaacactttt ttaaaggaca gttgaaaagg gcaagaggaa accagggcag 4260

ttctagagga gtgctggtga ctggatagca gttttaagtg gcgttcacct agtcaacacg 4320

accgcgtgtg ttgcccctgc cctgggctcc ccgccatgac atcttcacct tgcagcttgt 4380

gctgagactg acccaagtgc agctagcact gggacacaga tccttgtctt cagcaccttc 4440

caaggagcca acttttattc cctttcctct ctcccctccc cacctcgctt cttcccaatt 4500

tagtaactta gatgcttcca gcacatacgt aggtagctac cccagccggt ttggattaca 4560

ggcctgtgct ggaacatcat ctcagttggc caccttcctg gcaggctgta gacctgacat 4620

tttgagacaa gcctagagtc aggagcaggg actttgactc ttaggaagag cacacatgag 4680

ggcaaggctg ctggcagacg tctccattgt ccttatgttg tctgtgttgt attttttttt 4740

ttttattgac catggtgatt attttttaa accatcgtta atatactgaa gtgagctata 4800

gcacatatca tgtgcttagt ttgtttattt ttctccatct ccccttggct tcctagagtt 4860

tggacatatt ccaggctaaa tgcttttact caagactaca gaaaggtttg aagtagtgtg 4920

tgcatggcat gcacgtatgt aagtaatctg gggaagaagc aaagatctgt ttcattctta 4980

gcctcaggcc tcatgagggt ctccacaggg ccggagctca ggttacacca ctccttcgtc 5040

cttacaggag atgtagggag aagaatctgc aggctgcttg taggactgtt caccaagggg 5100

gataccagca gcaagagagt gcacccgttt agccctggac cctgtttctt actgtgtgac 5160

ttggctagag ttgggagttc ccccaaaata aacgtgtccc cattttacca gaaccaaacc 5220

tcaacacagc gaagctgtac tgtctttgtg tggcaaagat gttcccttgt aggccccttt 5280

caggtaaccg tcttcacaat gtattttcat cacagtttaa ggagcatcag ccgcttctca 5340

agtgggtagg gaaagcagaa aaacgtacgc aagaggacat ggatccaaaa tgatgatgaa 5400

gcatctccca tggggaggtg atggtgggga gatgatgggc taaacaggca acttttcaaa 5460

aacacagcta tcatagaaaa gaaacttgcc tcatgtaaac tggattgaga aattctcagt 5520

gattctgcaa tggatttttt tttaatgcag aagtaatgta tactctagta ttctggtgtt 5580

tttatattta tgtaataatt tcttaaaacc attcagacag ataactattt aatttttttt 5640

aagaaagttg gaaaggtctc tcctcccaag gacagtggct ggaagagttg gggcacagcc 5700

agttctgaat gttggtggag ggtgtagtgg ctttttggct cagcatccag aaacaccaaa 5760

ccaggctggc taaacaagtg gccgcgtgta aaaacagaca gctctgagtc aaatctgggc 5820

ccttccacaa gggtcctctg aaccaagccc cactcccttg ctaggggtga aagcattaca 5880

gagagatgga gccatctatc caagaagcct tcactcacct tcactgctgc tgttgcaact 5940

cggctgttct ggactctgat gtgtgtggag ggatggggaa tagaacattg actgtgttga 6000

ttaccttcac tattcggcca gcctgacctt ttaataactt tgtaaaaagc atgtatgtat 6060

ttatagtgtt ttagattttt ctaactttta tatcttaaaa gcagagcacc tgtttaagca 6120

ttgtacccct attgttaaag atttgtgtcc tctcattccc tctcttcctc ttgtaagtgc 6180

ccttctaata aacttttcat ggaaaagctc ctgtgccagg agctcagtct gaaaaaaaaa 6240

aaaaaaaaaa aaaaaa 6256

<210> 17
<211> 425
<212> PRT
<213> Homo sapiens

<400> 17

Met Ser Ser Ile Leu Pro Phe Thr Pro Pro Ile Val Lys Arg Leu Leu
1 5 10 15

Gly Trp Lys Lys Gly Glu Gln Asn Gly Gln Glu Glu Lys Trp Cys Glu
20 25 30

Lys Ala Val Lys Ser Leu Val Lys Lys Leu Lys Lys Thr Gly Gln Leu
35 40 45

Asp Glu Leu Glu Lys Ala Ile Thr Thr Gln Asn Val Asn Thr Lys Cys
50 55 60

Ile Thr Ile Pro Arg Ser Leu Asp Gly Arg Leu Gln Val Ser His Arg
65 70 75 80

Lys Gly Leu Pro His Val Ile Tyr Cys Arg Leu Trp Arg Trp Pro Asp
85 90 95

Leu His Ser His His Glu Leu Arg Ala Met Glu Leu Cys Glu Phe Ala
100 105 110

Phe Asn Met Lys Lys Asp Glu Val Cys Val Asn Pro Tyr His Tyr Gln
115 120 125

```
Arg Val Glu Thr Pro Val Leu Pro Pro Val Leu Val Pro Arg His Thr
    130                 135                 140

Glu Ile Pro Ala Glu Phe Pro Pro Leu Asp Asp Tyr Ser His Ser Ile
145                 150                 155                 160

Pro Glu Asn Thr Asn Phe Pro Ala Gly Ile Glu Pro Gln Ser Asn Ile
                165                 170                 175

Pro Glu Thr Pro Pro Pro Gly Tyr Leu Ser Glu Asp Gly Glu Thr Ser
            180                 185                 190

Asp His Gln Met Asn His Ser Met Asp Ala Gly Ser Pro Asn Leu Ser
        195                 200                 205

Pro Asn Pro Met Ser Pro Ala His Asn Asn Leu Asp Leu Gln Pro Val
    210                 215                 220

Thr Tyr Cys Glu Pro Ala Phe Trp Cys Ser Ile Ser Tyr Tyr Glu Leu
225                 230                 235                 240

Asn Gln Arg Val Gly Glu Thr Phe His Ala Ser Gln Pro Ser Met Thr
                245                 250                 255

Val Asp Gly Phe Thr Asp Pro Ser Asn Ser Glu Arg Phe Cys Leu Gly
            260                 265                 270

Leu Leu Ser Asn Val Asn Arg Asn Ala Ala Val Glu Leu Thr Arg Arg
        275                 280                 285

His Ile Gly Arg Gly Val Arg Leu Tyr Tyr Ile Gly Gly Glu Val Phe
    290                 295                 300

Ala Glu Cys Leu Ser Asp Ser Ala Ile Phe Val Gln Ser Pro Asn Cys
305                 310                 315                 320

Asn Gln Arg Tyr Gly Trp His Pro Ala Thr Val Cys Lys Ile Pro Pro
                325                 330                 335

Gly Cys Asn Leu Lys Ile Phe Asn Asn Gln Glu Phe Ala Ala Leu Leu
            340                 345                 350

Ala Gln Ser Val Asn Gln Gly Phe Glu Ala Val Tyr Gln Leu Thr Arg
        355                 360                 365
```

Met Cys Thr Ile Arg Met Ser Phe Val Lys Gly Trp Gly Ala Glu Tyr
370 375 380

Arg Arg Gln Thr Val Thr Ser Thr Pro Cys Trp Ile Glu Leu His Leu
385 390 395 400

Asn Gly Pro Leu Gln Trp Leu Asp Lys Val Leu Thr Gln Met Gly Ser
405 410 415

Pro Ser Ile Arg Cys Ser Ser Val Ser
420 425

<210> 18
<211> 3221
<212> DNA
<213> Homo sapiens

<400> 18
caacaacacg gccctggtcg tcgtcgccgc tgcggtaacg gagcggtttg ggtggcggag 60

cctgcgttcg cgccttccca ctcccctcgc caccgcccga gcccaggtta tcctgaatac 120

atgtctaaca attttccttg caacgttagc tgttgttttt cactgtttcc aaaggatcaa 180

aattgcttca gaaattggag acatatttga tttaaaagga aaaacttgaa caaatggaca 240

atatgtctat tacgaataca ccaacaagta atgatgcctg tctgagcatt gtgcatagtt 300

tgatgtgcca tagacaaggt ggagagagtg aaacatttgc aaaaagagca attgaaagtt 360

tggtaaagaa gctgaaggag aaaaaagatg aattggattc tttaataaca gctataacta 420

caaatggagc tcatcctagt aaatgtgtta ccatacagag aacattggat gggaggcttc 480

aggtggctgg tcggaaagga tttcctcatg tgatctatgc ccgtctctgg aggtggcctg 540

atcttcacaa aaatgaacta aaacatgtta aatattgtca gtatgcgttt gacttaaaat 600

gtgatagtgt ctgtgtgaat ccatatcact acgaacgagt tgtatcacct ggaattgatc 660

tctcaggatt aacactgcag agtaatgctc catcaagtat gatggtgaag gatgaatatg 720

tgcatgactt tgagggacag ccatcgttgt ccactgaagg acattcaatt caaaccatcc 780

agcatccacc aagtaatcgt gcatcgacag agacatacag caccccagct ctgttagccc 840

catctgagtc taatgctacc agcactgcca actttcccaa cattcctgtg gcttccacaa 900

gtcagcctgc cagtatactg gggggcagcc atagtgaagg actgttgcag atagcatcag 960

ggcctcagcc aggacagcag cagaatggat ttactggtca gccagctact taccatcata 1020

acagcactac cacctggact ggaagtagga ctgcaccata cacacctaat ttgcctcacc 1080

accaaaacgg ccatcttcag caccacccgc ctatgccgcc ccatcccgga cattactggc 1140

ctgttcacaa tgagcttgca ttccagcctc ccatttccaa tcatcctgct cctgagtatt 1200

ggtgttccat tgcttacttt gaaatggatg ttcaggtagg agagacattt aaggttcctt 1260

caagctgccc tattgttact gttgatggat acgtggaccc ttctggagga gatcgctttt 1320

gtttgggtca actctccaat gtccacagga cagaagccat tgagagagca aggttgcaca 1380

taggcaaagg tgtgcagttg gaatgtaaag gtgaaggtga tgtttgggtc aggtgcctta 1440

gtgaccacgc ggtctttgta cagagttact acttagacag agaagctggg cgtgcacctg 1500

gagatgctgt tcataagatc tacccaagtg catatataaa ggtctttgat ttgcgtcagt 1560

gtcatcgaca gatgcagcag caggcggcta ctgcacaagc tgcagcagct gcccaggcag 1620

cagccgtggc aggaaacatc cctggcccag gatcagtagg tggaatagct ccagctatca 1680

gtctgtcagc tgctgctgga attggtgttg atgaccttcg tcgcttatgc atactcagga 1740

tgagttttgt gaaaggctgg ggaccggatt acccaagaca gagcatcaaa gaaacacctt 1800

gctggattga aattcactta caccgggccc tccagctcct agacgaagta cttcatacca 1860

tgccgattgc agacccacaa cctttagact gaggtctttt accgttgggg cccttaacct 1920

tatcaggatg gtggactaca aaatacaatc ctgtttataa tctgaagata tatttcactt 1980

ttgttctgct ttatcttttc ataaagggtt gaaaatgtgt ttgctgcctt gctcctagca 2040

gacagaaact ggattaaaac aatttttttt ttcctcttca gaacttgtca ggcatggctc 2100

agagcttgaa gattaggaga aacacattct tattaattct tcacctgtta tgtatgaagg 2160

aatcattcca gtgctagaaa atttagccct ttaaaacgtc ttagagcctt ttatctgcag 2220

aacatcgata tgtatatcat tctacagaat aatccagtat tgctgatttt aaaggcagag 2280

aagttctcaa agttaattca cctatgttat tttgtgtaca agttgttatt gttgaacata 2340

cttcaaaaat aatgtgccat gtgggtgagt taattttacc aagagtaact ttactctgtg 2400

tttaaaaagt aagttaataa tgtattgtaa tctttcatcc aaaatatttt ttgcaagtta 2460

tattagtgaa gatggtttca attcagattg tcttgcaact tcagttttat ttttgccaag 2520

gcaaaaaact cttaatctgt gtgtatattg agaatccctt aaaattacca gacaaaaaaa 2580

tttaaaatta cgtttgttat tcctagtgga tgactgttga tgaagtatac ttttcccctg 2640

ttaaacagta gttgtattct tctgtatttc taggcacaag gttggttgct aagaagccta 2700

taagaggaat ttcttttcct tcattcatag ggaaaggttt tgtatttttt aaaacactaa 2760

aagcagcgtc actctaccta atgtctcact gttctgcaaa ggtggcaatg cttaaactaa 2820

ataatgaata aactgaatat tttggaaact gctaaattct atgttaaata ctgtgcagaa 2880

taatggaaac attacagttc ataataggta gtttggatat ttttgtactt gatttgatgt 2940

gacttttttt ggtataatgt ttaaatcatg tatgttatga tattgtttaa aattcagttt 3000

ttgtatcttg gggcaagact gcaaactttt ttatatcttt tggttattct aagccctttg 3060

ccatcaatga tcatatcaat tggcagtgac tttgtataga gaatttaagt agaaaagttg 3120

cagatgtatt gactgtacca cagacacaat atgtatgctt tttacctagc tggtagcata 3180

aataaaactg aatctcaaca taaaaaaaaa aaaaaaaaaa a 3221

<210> 19
<211> 552
<212> PRT
<213> Homo sapiens

<400> 19

Met Asp Asn Met Ser Ile Thr Asn Thr Pro Thr Ser Asn Asp Ala Cys
1 5 10 15

Leu Ser Ile Val His Ser Leu Met Cys His Arg Gln Gly Gly Glu Ser
20 25 30

Glu Thr Phe Ala Lys Arg Ala Ile Glu Ser Leu Val Lys Lys Leu Lys
35 40 45

Glu Lys Lys Asp Glu Leu Asp Ser Leu Ile Thr Ala Ile Thr Thr Asn
50 55 60

Gly Ala His Pro Ser Lys Cys Val Thr Ile Gln Arg Thr Leu Asp Gly
65 70 75 80

Arg Leu Gln Val Ala Gly Arg Lys Gly Phe Pro His Val Ile Tyr Ala
85 90 95

Arg Leu Trp Arg Trp Pro Asp Leu His Lys Asn Glu Leu Lys His Val
100 105 110

Lys Tyr Cys Gln Tyr Ala Phe Asp Leu Lys Cys Asp Ser Val Cys Val
115 120 125

Asn Pro Tyr His Tyr Glu Arg Val Val Ser Pro Gly Ile Asp Leu Ser
130 135 140

Gly Leu Thr Leu Gln Ser Asn Ala Pro Ser Ser Met Met Val Lys Asp
145 150 155 160

Glu Tyr Val His Asp Phe Glu Gly Gln Pro Ser Leu Ser Thr Glu Gly

```
                165                 170                 175

His Ser Ile Gln Thr Ile Gln His Pro Pro Ser Asn Arg Ala Ser Thr
            180                 185                 190

Glu Thr Tyr Ser Thr Pro Ala Leu Leu Ala Pro Ser Glu Ser Asn Ala
        195                 200                 205

Thr Ser Thr Ala Asn Phe Pro Asn Ile Pro Val Ala Ser Thr Ser Gln
    210                 215                 220

Pro Ala Ser Ile Leu Gly Gly Ser His Ser Glu Gly Leu Leu Gln Ile
225                 230                 235                 240

Ala Ser Gly Pro Gln Pro Gly Gln Gln Gln Asn Gly Phe Thr Gly Gln
                245                 250                 255

Pro Ala Thr Tyr His His Asn Ser Thr Thr Thr Trp Thr Gly Ser Arg
            260                 265                 270

Thr Ala Pro Tyr Thr Pro Asn Leu Pro His His Gln Asn Gly His Leu
        275                 280                 285

Gln His His Pro Pro Met Pro Pro His Pro Gly His Tyr Trp Pro Val
    290                 295                 300

His Asn Glu Leu Ala Phe Gln Pro Pro Ile Ser Asn His Pro Ala Pro
305                 310                 315                 320

Glu Tyr Trp Cys Ser Ile Ala Tyr Phe Glu Met Asp Val Gln Val Gly
                325                 330                 335

Glu Thr Phe Lys Val Pro Ser Ser Cys Pro Ile Val Thr Val Asp Gly
            340                 345                 350

Tyr Val Asp Pro Ser Gly Gly Asp Arg Phe Cys Leu Gly Gln Leu Ser
        355                 360                 365

Asn Val His Arg Thr Glu Ala Ile Glu Arg Ala Arg Leu His Ile Gly
    370                 375                 380

Lys Gly Val Gln Leu Glu Cys Lys Gly Glu Gly Asp Val Trp Val Arg
385                 390                 395                 400

Cys Leu Ser Asp His Ala Val Phe Val Gln Ser Tyr Tyr Leu Asp Arg
```

405 410 415

Glu Ala Gly Arg Ala Pro Gly Asp Ala Val His Lys Ile Tyr Pro Ser
420 425 430

Ala Tyr Ile Lys Val Phe Asp Leu Arg Gln Cys His Arg Gln Met Gln
435 440 445

Gln Gln Ala Ala Thr Ala Gln Ala Ala Ala Ala Ala Gln Ala Ala Ala
450 455 460

Val Ala Gly Asn Ile Pro Gly Pro Gly Ser Val Gly Gly Ile Ala Pro
465 470 475 480

Ala Ile Ser Leu Ser Ala Ala Ala Gly Ile Gly Val Asp Asp Leu Arg
485 490 495

Arg Leu Cys Ile Leu Arg Met Ser Phe Val Lys Gly Trp Gly Pro Asp
500 505 510

Tyr Pro Arg Gln Ser Ile Lys Glu Thr Pro Cys Trp Ile Glu Ile His
515 520 525

Leu His Arg Ala Leu Gln Leu Leu Asp Glu Val Leu His Thr Met Pro
530 535 540

Ile Ala Asp Pro Gln Pro Leu Asp
545 550

<210> 20
<211> 6504
<212> DNA
<213> Mus musculus

<400> 20
gagagcgcgc ctgggcggcg gggaggactt gcactaagaa gaagcctatg gcacctgtca 60

agttaaatgt cactccccgc ctccacttgg actttctgct taagacctgc atgtgacttt 120

tcacctgcga gccacgcttt tggtatctac tgactttgat tacaggaaag tgtctgaaga 180

tttgtatcaa atgacgtcaa tggccagctt gttttctttc actagtccag ccgtgaagcg 240

attgttgggc tggaaacaag gtgacgagga agagaaatgg gcagaaaagg cagtggatgc 300

tttagtgaaa aagctgaaga agaagaaggg tgctatggag gagctggaga aagccttgag 360

cagcccagga cagccaagca agtgtgtcac gatccccagg tccttggatg gacgtctgca 420

agtttctcac aggaaaggct tgccccatgt tatatattgc cgtgtttggc gctggccaga 480

tttgcagagc catcacgagc taaaaccatt ggatatttgt gaatttcctt ttggatctaa 540

gcaaaaggaa gtttgtatca atccatacca ctataagaga gtggagagtc cagtcttacc 600

tccagtatta gtgcctcgtc acaatgaatt caatccacaa cacagccttc tggttcagtt 660

caggaacctg agccacaatg aaccgcacat gccacaaaac gccacgtttc ccgattcttt 720

ccaccaaccc aacaacgctc ctttcccctt atctcctaac agcccctatc ctccttcccc 780

tgctagcagc acatatccca actccccagc aagctctgga cctggaagtc catttcaact 840

cccagctgac acccctcccc ctgcctatat gccacctgat gatcagatgg ccccagataa 900

ttcccagcct atggatacaa gcagtaacat gattcctcag accatgccca gcatatccag 960

cagagatgtt cagcctgtcg cctatgagga gcccaaacac tggtgttcga ttgtctacta 1020

tgaattaaac aatcgtgttg gggaagcttt tcatgcatct tctactagtg tgttagtaga 1080

tggatttaca gatccttcaa ataacaaaag tagattctgc ctgggattgt tgtcaaatgt 1140

taatcgtaat tcaactattg aaaacactag gcggcatatt ggaaaaggtg ttcatctata 1200

ctacgttggt ggggaggtgt acgctgagtg tcttagtgac agcagcatct ttgttcagag 1260

taggaactgc aactttcacc atggcttcca tcccaccacc gtctgtaaga tccccagcag 1320

ctgcagcctc aagattttta acaatcagga gtttgctcag cttctggctc agtcagtcaa 1380

ccatggattc gaggctgtgt atgagctcac caagatgtgt accattcgaa tgagctttgt 1440

caagggctgg ggagcagagt accaccgaca ggacgtcacc agtactccct gctggattga 1500

gattcacctc cacgggcctc tgcagtggct ggataaagtc cttactcaga tgggctctcc 1560

gctgaacccc atttcttctg tttcatagtg cagaagtatt ctttcaacta tatttttagt 1620

ggacttgttt taattttaga ggaatttcca gtacagatgc tgtgagctga catggaaaac 1680

agatattatt ttttctacgt aattgtgacc aacacatttg tattttatga tgatattaca 1740

tttgtttgta ttcgtgttca ttgtgattaa ctttcaaaag tattgtaaac gatgtagagt 1800

attttgcccc tgttgaaatg tttagcattg atcttaaact ggaacgtact ttttcttatt 1860

gtcccaacgt tttttaattt gttaaatttt ttttacaaag tagttcatca cataatgaaa 1920

ttttatccta taagagaaca tatattgtgg aaagcagtag atgatatttc tctgggaatt 1980

tctttgcctt accacctttg aaaaagcata cattgtttgc aaaacctcaa agtagggctt 2040

gcttaaagga aactgttgaa tcttgtttga aggacactgc agtcctaacg tgttcagtga 2100

aagcaaggtg gtagatttct ggacgtcata catttacatt taatataggt aatattcatc 2160

agtgtaatgt gacttcatgc catatatatt ttgtaaaaca attccttttt aaaaacttca 2220

agtatttctc atttactcaa atttgttgta agtcctactt aacagttagt tactatgtgc 2280

tctgtggcct tgttcagcat tgtttgctgc tttgggccaa caattcaaga actctaattt 2340

tcctgtgcat taatcttttc attttgcact tttatgggtg actgtcttag tgtagcctct 2400

ggtaaaatac tattaggtgg cctggtttta gagctcctcc tcgctgcctt ggcactcctt 2460

tgtgcaacac gaccacttag agatgacagc tgtgagctgt gctgcttttt ctagccttta 2520

atttccaatg tagtttataa tgttgttctt ctatagctcc agctaaggtg cctgttagtc 2580

ccctacaatg ttatgagcat tattgacatt gaaaggttat gtatgtatga atacctttgc 2640

tccttaccag acttgtcata caaggactcg tgcagtgtag ccagtagagg ctctttggtt 2700

ggcccaagaa tgaggctgtt ggtgtaagtg aatcacaata gggattggga tagttcatgt 2760

catatgtcat atagcaagac aatgtagagt gtaggcttgt ctctctgcat caacgctctg 2820

cctctttctt tttatccttt tagaacctac atggacgcta atctccacaa cactgttgga 2880

tgtgaacact cttaagacac tcatctagtt cactgtgcct tgtccttagg actcttaacc 2940

actttctagg gagcagttat ggcctgagat ggacagtcat ggcctgagaa tgaagacact 3000

actttgataa agaaaaaggc ctcatttgcc tatcagagtg agaaaggttt ttttctggtg 3060

ccttttgaaa atatacagag ccacttggtt cttctgctga aaatgtaatt ttggtttgac 3120

tttttagagt gcccttcctg cctttatgag gaaaacagct atttttttt ttgggggggg 3180

ggattccttt tgttttctgt gccattattt attgcctttc agagtgcaac cattgggtgg 3240

ctttgctcct tcagagaggg ctccttgata gccttcagta gcttgagctg tagacataag 3300

tattccatag caagagtgtg tcagctccat gagagagatg tctgctttat agccgaggca 3360

gaaaccgttc atgttccttt acttggcagc cttcaggaac aggtttgtaa gaacgtgtct 3420

tgagttgagt gagtgtatgt ctgtgagctc tgctgaagtc tggacacaag ggccttgcct 3480

gctccttttt tcagcagtgg gttacatgtt gtctctccac agtcttcatg tcataggtct 3540

cggacttgca gagtcctatg tggcctgcca tctgtacagt ggcaggactg aagctctgag 3600

ctgttctgag gttcatggag aaatcccaac ctattctgtg gtcagtaaat ggagactgtg 3660

tagtctacct gctcctgtac tgtccttact gtatgtaagg atatacagac gcctgtgggt 3720

aggcagtact cacagtgaga tgaagacagc aagtgtgcac tgaaccacag agggcaggga 3780

gtagggcctc tgaagaagcc accagaccag accagtgccg gtacagtctt tgtcagagat 3840

ggctctgatg gggcccagac tgaccctgac catgctgagt tgctgagggt agccttcagt 3900

tctctaccct ctgaagtgct aggatgacag acatccgcca tcatacccag cttccgtggt 3960

gctaaggatc agcctcagtc ttcaggcgtg ctaggcatgt actttgccaa gtatttagta 4020

tacaaaatac attagtatct gccagggaaa aaagatttgc aaataataaa gattgccatc 4080

agtttgataa atgttgtaaa tggaagaatc aaaatctcag cgatggatta cagcaacaag 4140

atgctgccta ggaaaagcag gaccaagagg tacatttgac tagtatacct tcagcgtagc 4200

gtgatgacct cactgatgtc acccaactga acttaagggc tgtaagtagg cgtgctgtgg 4260

gccttccaga actagagaaa attataggag gaagtcagtt ctaaagtatc aaaagctggg 4320

taatggtggc acatgccttt gattctagca ctcgggaagc aggggctagc ctagtctaca 4380

gagcaacttc tacacagaga aactgtcttg gaggaaaata aaaaaagaaa agtcaaagag 4440

caaacaaata gaacagagta ggaatccgtg tccccttttt tctatgtttc acggttgcag 4500

gtgtaagaaa agtagtcata gatgtggctg agtttctaag atgaaaccag tagtaagatt 4560

gctaaatata acacttcaac caagttaaac acccttgggg ggtatgaatg aaagtaacac 4620

tgcaatatga aatgaaccgt gcaagtaaca cttggggtta cctcacagtc tccctatgcc 4680

tgagaggact gtgggaaaca tttccatccc ctgccagtat cgccattggg aggacagagt 4740

agatgaagaa gtgaagtctt actggtccag ggcacgcctg tcagcaatgc catttgtgct 4800

tctgccacag agagcaccga gaggcttggc tcagtatcct cgaaccttct ctggtcactt 4860

ccctggcagc acttgggtcc ctgtcactca ctggtctctt aaaagtcccg tctctttgct 4920

tcctaaagat tctctaaaaa aattactatt ttttatttct tttttaaaag tctttgttat 4980

tttgttttgg gatacagtct ctttgtacag tcctggctgg cctggaaatt actatgtagg 5040

ccagcctcaa acttgaagta attctcttgc ctctgcctct ggagttctgg gattacaggc 5100

atgcactgca gagtacagtg agctctgatg gcttttaaaa ttcagcccct ttgagggttt 5160

ggttttagat ccattagctt tgtctgaacc catctttgtc cggccgagta aatcctctgc 5220

tatccggggt ctcggtagaa atgtgttctc agtatacata cgactaaaca ttggttgttt 5280

ataggtagcc tcagatattt ggtagagcat ctttttttgaa agtaatctcc agctaggtgg 5340

gtatttccct cacagcagta ggattttccc tttaggagat accagttctt catctttctt 5400

gtgaaaataa tgcctttatg gggagtgaag attaaggagt tgtttctaca ctaacagaat 5460

tctatttgat ggacaacttg gacagttctg tggacttggg tgggttctag tgtgctaaga 5520

aggataacag tatttaatag tgtctgtcat caggccttgc tcatctccct gtctagggct 5580

gtaggtcagt gctcgagcac ttagcaggca tcgagtctag tgttcagtgc ccagcattgc 5640

acagaactca gaatatatct gtactgaaac tgaagtgacc acctacaacc aggtggtatg 5700

ccagaaccac agaaaggaga ttcacggtga tgtgtttaaa gcattgggct ggtgacggtt 5760

gctgtgtagt aatgacctct tcctcagcaa agagagtcct ggagcaggct gtcctcagaa 5820

gagggaaggg actggtgtgc tccttgtgca gataacttag tgtataaatc ggcatgagta 5880

gctatccttt aaggatttgt ttgaagttac tctttgtaaa aagttgagaa ttttgtgtgc 5940

agttgggcac atgcttgccc ttcccccacc cgccatagtc ctgcctctct tgctgtgaac 6000

tggtgtcagc tacaacactc cagctaggtc tgagctcttt tgagagaagg tctcgtagag 6060

caccattctc agagagaagc taaagcatgg ggagccttag gacggtcagg caatgcactc 6120

tttaccacgg ctggctaagg ctgcagcttg accgtcctta cctaaatcag gtaagaatgt 6180

gattacagag cgagtgcttg tgttccccgg cctgccttct ccgaggaaga tgcttcatcc 6240

gaggatgatg cagagcagac gatggctgca ctgtaggtct gcctccttct gtgtatgggt 6300

tctgctgctg cttacggcat aggaaagtac actagcagcg tgcttcaatt ctgccatctt 6360

ttgatactta taaaaatgta ttaggtttta ctgtattgtg ctctcaaagc cataactctt 6420

aagaaatttg gttttttttgc atattgtttg ctaatacttt gttttaataa acctcaaaat 6480

ctgcttaaaa aaaaaaaaaa aaaa 6504

<210> 21
<211> 465
<212> PRT
<213> Mus musculus

<400> 21

Met Thr Ser Met Ala Ser Leu Phe Ser Phe Thr Ser Pro Ala Val Lys
1 5 10 15

Arg Leu Leu Gly Trp Lys Gln Gly Asp Glu Glu Glu Lys Trp Ala Glu
20 25 30

Lys Ala Val Asp Ala Leu Val Lys Lys Leu Lys Lys Lys Lys Gly Ala
35 40 45

Met Glu Glu Leu Glu Lys Ala Leu Ser Ser Pro Gly Gln Pro Ser Lys
50 55 60

Cys Val Thr Ile Pro Arg Ser Leu Asp Gly Arg Leu Gln Val Ser His
65 70 75 80

Arg Lys Gly Leu Pro His Val Ile Tyr Cys Arg Val Trp Arg Trp Pro
85 90 95

Asp Leu Gln Ser His His Glu Leu Lys Pro Leu Asp Ile Cys Glu Phe
100 105 110

Pro Phe Gly Ser Lys Gln Lys Glu Val Cys Ile Asn Pro Tyr His Tyr

115 120 125

Lys Arg Val Glu Ser Pro Val Leu Pro Pro Val Leu Val Pro Arg His
130 135 140

Asn Glu Phe Asn Pro Gln His Ser Leu Leu Val Gln Phe Arg Asn Leu
145 150 155 160

Ser His Asn Glu Pro His Met Pro Gln Asn Ala Thr Phe Pro Asp Ser
165 170 175

Phe His Gln Pro Asn Asn Ala Pro Phe Pro Leu Ser Pro Asn Ser Pro
180 185 190

Tyr Pro Pro Ser Pro Ala Ser Ser Thr Tyr Pro Asn Ser Pro Ala Ser
195 200 205

Ser Gly Pro Gly Ser Pro Phe Gln Leu Pro Ala Asp Thr Pro Pro Pro
210 215 220

Ala Tyr Met Pro Pro Asp Asp Gln Met Ala Pro Asp Asn Ser Gln Pro
225 230 235 240

Met Asp Thr Ser Ser Asn Met Ile Pro Gln Thr Met Pro Ser Ile Ser
245 250 255

Ser Arg Asp Val Gln Pro Val Ala Tyr Glu Glu Pro Lys His Trp Cys
260 265 270

Ser Ile Val Tyr Tyr Glu Leu Asn Asn Arg Val Gly Glu Ala Phe His
275 280 285

Ala Ser Ser Thr Ser Val Leu Val Asp Gly Phe Thr Asp Pro Ser Asn
290 295 300

Asn Lys Ser Arg Phe Cys Leu Gly Leu Leu Ser Asn Val Asn Arg Asn
305 310 315 320

Ser Thr Ile Glu Asn Thr Arg Arg His Ile Gly Lys Gly Val His Leu
325 330 335

Tyr Tyr Val Gly Gly Glu Val Tyr Ala Glu Cys Leu Ser Asp Ser Ser
340 345 350

Ile Phe Val Gln Ser Arg Asn Cys Asn Phe His His Gly Phe His Pro

355 360 365

Thr Thr Val Cys Lys Ile Pro Ser Ser Cys Ser Leu Lys Ile Phe Asn
370 375 380

Asn Gln Glu Phe Ala Gln Leu Leu Ala Gln Ser Val Asn His Gly Phe
385 390 395 400

Glu Ala Val Tyr Glu Leu Thr Lys Met Cys Thr Ile Arg Met Ser Phe
405 410 415

Val Lys Gly Trp Gly Ala Glu Tyr His Arg Gln Asp Val Thr Ser Thr
420 425 430

Pro Cys Trp Ile Glu Ile His Leu His Gly Pro Leu Gln Trp Leu Asp
435 440 445

Lys Val Leu Thr Gln Met Gly Ser Pro Leu Asn Pro Ile Ser Ser Val
450 455 460

Ser
465

<210> 22
<211> 2887
<212> DNA
<213> Homo sapiens

<400> 22
gaattcggca cgaggagcga tcgagggagc tgagccgaga gaaagagccg ccgggcgctg 60

cctcgccaga cctcgctggg accccggggc caccgggagg cacttttgtg gaggggggag 120

gggggggcgac ctcggcagcc tcggcgcacg aagcgtccga gggcagcgtg gggygggctg 180

cgacctctgc atcggtggac tgcattttta attaaggatt cccagcagct ctttgggatt 240

tttacagctt ccactcatgt gttgacaccc gcgtccagga gaaactcgct ccaagtgcat 300

ctagcgcctg ggacctgaga cggcgttggc ctttcgtgca tgcaaatcca gggatttagg 360

ttttgtttgg gatttccttt tctttctttc ctttttttt tctttttgca gggagtaaga 420

agggagctgg gggtatcaac aagcctgcct ttcggatcct gcgggaaaag cccatgtagt 480

taagcgcttt ggtttaaaaa aaaggcaagg taaaggcagg gctttccaga cacatttagg 540

ggttcgcgcg agctttgtgc tcatggacca gccgcacaac ttttgaaggc tcgccggccc 600

atgtggggtc tttctggcgg gccgcctgca gccccctaa agcgcggggg gctggagttg 660

ttgagcagcc ccgccgctgt ggtccatgta gccgctggcc gcgcgcggaa tgcggcttcg 720

cgtgcgcgtg ttcccggccg tcccgccttg gcgaggtccc tcatgttgtt gccctgcggc 780

gccccttga agacaggctg tgcgcggttt gcacggcggt ccgcggcgga ggttcatgtg 840

ggggtgcgac ccgcgcagcc ggcgccttgt tgagggaacg gacccccggt aaccggagac 900

cgcctccccc ccacccctgg cgccaaagga tattgtatgt tcaggtccaa acgctcgggg 960

ctggtgcggc gactttggcg aagtcgtgtg gtccccaacc gggaggaagg cggcagcggc 1020

ggcggcggtg gcggcgacga ggatgggagc ttgggcagcc gagctgagcc ggccccgcgg 1080

gcaagagagg gcggaggctg cggccgctcc gaagtccgcc cggtagcccc gcggcggccc 1140

cgggacgcag tgggacagcg aggcgcccag ggcgcgggga ggcgccggcg cgcaggggc 1200

cccccgaggc ccatgtcgga gccaggggcc ggcgctggga gctccctgct ggacgtggcg 1260

gagccgggag gcccgggctg gctgcccgag agtgactgcg agacggtgac ctgctgtctc 1320

ttttcggagc gggacgccgc cggcgcgccc cgggacgcca gcgacccccct ggccggggcg 1380

gccctggagc cggcgggcgg cgggcggagt cgcgaagcgc gctcgcggct gctgctgctg 1440

gagcaggaac tcaaaaccgt cacgtactcg ctgctgaagc ggctcaagga gcgctcgctg 1500

gacacgctgc tggaggcggt ggagtcccgc ggcggcgtgc cgggcggctg cgtgctggtg 1560

ccgcgcgccg acctccgcct gggcggccag cccgcgccgc cgcagctgct gctcggccgc 1620

ctctttcgct ggcccgacct gcagcacgcc gtggagctga agcccctgtg cggctgccac 1680

agcttcgccg ccgccgccga cggccctacc gtgtgctgca acccctacca cttcagccgg 1740

ctctgcgggc ccgaatctcc gccacctccc tactctcggc tgtctcctcg cgacgagtac 1800

aagccactgg atctgtccga ttccacattg tcttacactg aaacggaggc taccaactcc 1860

ctcatcactg ctccgggtga attctcagac gccagcatgt ctccggacgc caccaagccg 1920

agccactggt gcagcgtggc gtactgggag caccggacgc gcgtgggccg cctctatgcg 1980

gtgtacgacc aggccgtcag catcttctac gacctacctc agggcagcgg cttctgcctg 2040

ggccagctca acctggagca gcgcagcgag tcggtgcggc gaacgcgcag caagatcggc 2100

ttcggcatcc tgctcagcaa ggagcccgac ggcgtgtggg cctacaaccg cggcgagcac 2160

cccatcttcg tcaactcccc gacgctggac gcgcccggcg gccgcgccct ggtcgtgcgc 2220

aaggtgcccc ccggctactc catcaaggtg ttcgacttcg agcgctcggg cctgcagcac 2280

gcgcccgagc ccgacgccgc cgacggcccc tacgacccca acagcgtccg catcagcttc 2340

gccaagggct gggggccctg ctactcccgg cagttcatca cctcctgccc ctgctggctg 2400

gagatcctcc tcaacaaccc cagatagtgg cggccccggc gggaggggcg ggtgggaggc 2460

cgcggccacc gccacctgcc ggcctcgaga ggggccgatg cccagagaca cagcccccac 2520

ggacaaaacc ccccagatat catctaccta gatttaatat aaagttttat atattatatg 2580

gaaatatata ttatacttgt aattatggag tcatttttac aatgtaatta tttatgtatg 2640

gtgcaatgtg tgtatatgga caaaacaaga aagacgcact ttggcttata attctttcaa 2700

tacagatata ttttctttct cttcctcctt cctcttcctt acttttttata tatatatata 2760

aagaaaatga tacagcagag ctaggtggaa aagcctgggt ttggtgtatg gtttttgaga 2820

tattaatgcc cagacaaaaa gctaatacca gtcactcgat aataaagtat tcgcattata 2880

aaaaaga 2887

<210> 23
<211> 496
<212> PRT
<213> Homo sapiens

<400> 23

Met Phe Arg Ser Lys Arg Ser Gly Leu Val Arg Arg Leu Trp Arg Ser
1 5 10 15

Arg Val Val Pro Asn Arg Glu Glu Gly Gly Ser Gly Gly Gly Gly Gly
20 25 30

Gly Asp Glu Asp Gly Ser Leu Gly Ser Arg Ala Glu Pro Ala Pro Arg
35 40 45

Ala Arg Glu Gly Gly Gly Cys Gly Arg Ser Glu Val Arg Pro Val Ala
50 55 60

Pro Arg Arg Pro Arg Asp Ala Val Gly Gln Arg Gly Ala Gln Gly Ala
65 70 75 80

Gly Arg Arg Arg Arg Ala Gly Gly Pro Pro Arg Pro Met Ser Glu Pro
85 90 95

Gly Ala Gly Ala Gly Ser Ser Leu Leu Asp Val Ala Glu Pro Gly Gly
100 105 110

Pro Gly Trp Leu Pro Glu Ser Asp Cys Glu Thr Val Thr Cys Cys Leu
115 120 125

Phe Ser Glu Arg Asp Ala Ala Gly Ala Pro Arg Asp Ala Ser Asp Pro
130 135 140

Leu Ala Gly Ala Ala Leu Glu Pro Ala Gly Gly Gly Arg Ser Arg Glu

145 150 155 160

Ala Arg Ser Arg Leu Leu Leu Leu Glu Gln Glu Leu Lys Thr Val Thr
165 170 175

Tyr Ser Leu Leu Lys Arg Leu Lys Glu Arg Ser Leu Asp Thr Leu Leu
180 185 190

Glu Ala Val Glu Ser Arg Gly Gly Val Pro Gly Gly Cys Val Leu Val
195 200 205

Pro Arg Ala Asp Leu Arg Leu Gly Gly Gln Pro Ala Pro Pro Gln Leu
210 215 220

Leu Leu Gly Arg Leu Phe Arg Trp Pro Asp Leu Gln His Ala Val Glu
225 230 235 240

Leu Lys Pro Leu Cys Gly Cys His Ser Phe Ala Ala Ala Ala Asp Gly
245 250 255

Pro Thr Val Cys Cys Asn Pro Tyr His Phe Ser Arg Leu Cys Gly Pro
260 265 270

Glu Ser Pro Pro Pro Pro Tyr Ser Arg Leu Ser Pro Arg Asp Glu Tyr
275 280 285

Lys Pro Leu Asp Leu Ser Asp Ser Thr Leu Ser Tyr Thr Glu Thr Glu
290 295 300

Ala Thr Asn Ser Leu Ile Thr Ala Pro Gly Glu Phe Ser Asp Ala Ser
305 310 315 320

Met Ser Pro Asp Ala Thr Lys Pro Ser His Trp Cys Ser Val Ala Tyr
325 330 335

Trp Glu His Arg Thr Arg Val Gly Arg Leu Tyr Ala Val Tyr Asp Gln
340 345 350

Ala Val Ser Ile Phe Tyr Asp Leu Pro Gln Gly Ser Gly Phe Cys Leu
355 360 365

Gly Gln Leu Asn Leu Glu Gln Arg Ser Glu Ser Val Arg Arg Thr Arg
370 375 380

Ser Lys Ile Gly Phe Gly Ile Leu Leu Ser Lys Glu Pro Asp Gly Val

385 390 395 400

Trp Ala Tyr Asn Arg Gly Glu His Pro Ile Phe Val Asn Ser Pro Thr
405 410 415

Leu Asp Ala Pro Gly Gly Arg Ala Leu Val Val Arg Lys Val Pro Pro
420 425 430

Gly Tyr Ser Ile Lys Val Phe Asp Phe Glu Arg Ser Gly Leu Gln His
435 440 445

Ala Pro Glu Pro Asp Ala Ala Asp Gly Pro Tyr Asp Pro Asn Ser Val
450 455 460

Arg Ile Ser Phe Ala Lys Gly Trp Gly Pro Cys Tyr Ser Arg Gln Phe
465 470 475 480

Ile Thr Ser Cys Pro Cys Trp Leu Glu Ile Leu Leu Asn Asn Pro Arg
485 490 495

<210> 24
<211> 3103
<212> DNA
<213> Homo sapiens

<400> 24
cggagagccg cgcagggcgc gggccgcgcg gggtggggca gccggagcgc aggcccccga 60

tccccggcgg gcgccccggg gccccgcgc gcgccccggc ctccgggaga ctggcgcatg 120

ccacggagcg cccctcgggc cgccgccgct cctgcccggg cccctgctgc tgctgctgtc 180

gcctgcgcct gctgccccaa ctcggcgccc gacttcttca tggtgtgcgg aggtcatgtt 240

cgctccttag caggcaaacg acttttctcc tcgcctcctc gccccgcatg ttcaggacca 300

aacgatctgc gctcgtccgg cgtctctgga ggagccgtgc gcccggcggc gaggacgagg 360

aggagggcgc agggggaggt ggaggaggag gcgagctgcg gggagaaggg gcgacggaca 420

gccgagcgca tggggccggt ggcggcggcc cgggcagggc tggatgctgc ctgggcaagg 480

cggtgcgagg tgccaaaggt caccaccatc cccacccgcc agccgcgggc gccggcgcgg 540

ccgggggcgc cgaggcggat ctgaaggcgc tcacgcactc ggtgctcaag aaactgaagg 600

agcggcagct ggagctgctg ctccaggccg tggagtcccg cggcgggacg cgcaccgcgt 660

gcctcctgct gcccggccgc ctggactgca ggctgggccc gggggcgccc gccggcgcgc 720

agcctgcgca gccgccctcg tcctactcgc tccccctcct gctgtgcaaa gtgttcaggt 780

ggccggatct caggcattcc tcggaagtca agaggctgtg ttgctgtgaa tcttacggga 840

agatcaaccc cgagctggtg tgctgcaacc cccatcacct tagccgactc tgcgaactag 900

agtctccccc ccctccttac tccagatacc cgatggattt tctcaaacca actgcagact 960

gtccagatgc tgtgccttcc tccgctgaaa caggggggaac gaattatctg gcccctgggg 1020

ggctttcaga ttcccaactt cttctggagc ctggggatcg gtcacactgg tgcgtggtgg 1080

catactggga ggagaagacg agagtgggga ggctctactg tgtccaggag ccctctctgg 1140

atatcttcta tgatctacct caggggaatg gcttttgcct cggacagctc aattcggaca 1200

acaagagtca gctggtgcag aaggtgcgga gcaaaatcgg ctgcggcatc cagctgacgc 1260

gggaggtgga tggtgtgtgg gtgtacaacc gcagcagtta ccccatcttc atcaagtccg 1320

ccacactgga caacccggac tccaggacgc tgttggtaca caaggtgttc cccggtttct 1380

ccatcaaggc tttcgactac gagaaggcgt acagcctgca gcggcccaat gaccacgagt 1440

ttatgcagca gccgtggacg ggctttaccg tgcagatcag ctttgtgaag ggctggggtc 1500

agtgctacac ccgccagttc atcagcagct gcccgtgctg gctagaggtc atcttcaaca 1560

gccggtagcc gcgtgcggag gggacagagc gtgagctgag caggccacac ttcaaactac 1620

tttgctgcta atattttcct cctgagtgct tgcttttcat gcaaactctt tggtcgtttt 1680

ttttttgttt gttggttggt tttcttcttc tcgtcctcgt ttgtgttctg ttttgtttcg 1740

ctctttgaga aatagcttat gaaaagaatt gttgggggtt tttttggaag aaggggcagg 1800

tatgatcggc aggacaccct gataggaaga ggggaagcag aaatccaagc accaccaaac 1860

acagtgtatg aaggggggcg gtcatcattt cacttgtcag gagtgtgtgt gagtgtgagt 1920

gtgcggctgt gtgtgcacgc gtgtgcagga gcggcagatg gggagacaac gtgctctttg 1980

ttttgtgtct cttatggatg tccccagcag agaggtttgc agtcccaagc ggtgtctctc 2040

ctgccccttg gacacgctca gtggggcaga ggcagtacct gggcaagctg gcggctgggg 2100

tcccagcagc tgccaggagc acggctctgt ccccagcctg ggaaagcccc tgcccctcct 2160

ctccctcatc aaggacacgg gcctgtccac aggcttctga gcagcgagcc tgctagtggc 2220

cgaaccagaa ccaattattt tcatccttgt cttattccct tcctgccagc ccctgccatt 2280

gtagcgtctt tctttttttgg ccatctgctc ctggatctcc ctgagatggg cttcccaagg 2340

gctgccgggg cagccccctc acagtattgc tcacccagtg ccctctcccc tcagcctctc 2400

ccctgcctgc cctggtgaca tcaggttttt cccggactta gaaaaccagc tcagcactgc 2460

ctgctcccat cctgtgtgtt aagctctgct attaggccag caagcgggga tgtccctggg 2520

agggacatgc ttagcagtcc ccttccctcc aagaaggatt tggtccgtca taacccaagg 2580

taccatccta ggctgacacc taactcttct ttcatttctt ctacaactca tacactcgta 2640

tgatacttcg acactgttct tagctcaatg agcatgttta gactttaaca taagctattt 2700

ttctaactac aaaggtttaa atgaacaaga gaagcattct cattggaaat ttagcattgt 2760

agtgctttga gagagaaagg actcctgaaa aaaaacctga gatttattaa agaaaaaaat 2820

gtattttatg ttatatataa atatattatt acttgtaaat ataaagacgt tttataagca 2880

tcattattta tgtattgtgc aatgtgtata aacaagaaaa ataaagaaaa gatgcacttt 2940

gctttaatat aaatgcaaat aacaaatgcc aaattaaaaa agataaacac aagattggtg 3000

ttttttccta tgggtgttat cacctagctg aatgtttttc taaaggagtt tatgttccat 3060

taaacgattt ttaaaatgta cacttgaaaa aaaaaaaaaa aaa 3103

<210> 25
<211> 426
<212> PRT
<213> Homo sapiens

<400> 25

Met Phe Arg Thr Lys Arg Ser Ala Leu Val Arg Arg Leu Trp Arg Ser
1 5 10 15

Arg Ala Pro Gly Gly Glu Asp Glu Glu Glu Gly Ala Gly Gly Gly Gly
20 25 30

Gly Gly Gly Glu Leu Arg Gly Glu Gly Ala Thr Asp Ser Arg Ala His
35 40 45

Gly Ala Gly Gly Gly Gly Pro Gly Arg Ala Gly Cys Cys Leu Gly Lys
50 55 60

Ala Val Arg Gly Ala Lys Gly His His His Pro His Pro Pro Ala Ala
65 70 75 80

Gly Ala Gly Ala Ala Gly Gly Ala Glu Ala Asp Leu Lys Ala Leu Thr
85 90 95

His Ser Val Leu Lys Lys Leu Lys Glu Arg Gln Leu Glu Leu Leu Leu
100 105 110

Gln Ala Val Glu Ser Arg Gly Gly Thr Arg Thr Ala Cys Leu Leu Leu
115 120 125

Pro Gly Arg Leu Asp Cys Arg Leu Gly Pro Gly Ala Pro Ala Gly Ala
130 135 140

```
Gln Pro Ala Gln Pro Pro Ser Ser Tyr Ser Leu Pro Leu Leu Leu Cys
145                 150                 155                 160

Lys Val Phe Arg Trp Pro Asp Leu Arg His Ser Ser Glu Val Lys Arg
                165                 170                 175

Leu Cys Cys Cys Glu Ser Tyr Gly Lys Ile Asn Pro Glu Leu Val Cys
            180                 185                 190

Cys Asn Pro His His Leu Ser Arg Leu Cys Glu Leu Glu Ser Pro Pro
        195                 200                 205

Pro Pro Tyr Ser Arg Tyr Pro Met Asp Phe Leu Lys Pro Thr Ala Asp
    210                 215                 220

Cys Pro Asp Ala Val Pro Ser Ser Ala Glu Thr Gly Gly Thr Asn Tyr
225                 230                 235                 240

Leu Ala Pro Gly Gly Leu Ser Asp Ser Gln Leu Leu Leu Glu Pro Gly
                245                 250                 255

Asp Arg Ser His Trp Cys Val Val Ala Tyr Trp Glu Glu Lys Thr Arg
            260                 265                 270

Val Gly Arg Leu Tyr Cys Val Gln Glu Pro Ser Leu Asp Ile Phe Tyr
        275                 280                 285

Asp Leu Pro Gln Gly Asn Gly Phe Cys Leu Gly Gln Leu Asn Ser Asp
    290                 295                 300

Asn Lys Ser Gln Leu Val Gln Lys Val Arg Ser Lys Ile Gly Cys Gly
305                 310                 315                 320

Ile Gln Leu Thr Arg Glu Val Asp Gly Val Trp Val Tyr Asn Arg Ser
                325                 330                 335

Ser Tyr Pro Ile Phe Ile Lys Ser Ala Thr Leu Asp Asn Pro Asp Ser
            340                 345                 350

Arg Thr Leu Leu Val His Lys Val Phe Pro Gly Phe Ser Ile Lys Ala
        355                 360                 365

Phe Asp Tyr Glu Lys Ala Tyr Ser Leu Gln Arg Pro Asn Asp His Glu
    370                 375                 380
```

Phe Met Gln Gln Pro Trp Thr Gly Phe Thr Val Gln Ile Ser Phe Val
385 390 395 400

Lys Gly Trp Gly Gln Cys Tyr Thr Arg Gln Phe Ile Ser Ser Cys Pro
405 410 415

Cys Trp Leu Glu Val Ile Phe Asn Ser Arg
420 425

<210> 26
<211> 1600
<212> DNA
<213> Mus musculus

<400> 26
tttgaatcgg accccactgc ttccagaagg aaaggcaagc ttgttcctat gacatccgtg 60

gacaggtact tgccgccgac ctgcccgggg ccctgcaagc cttgaaaggt ctcatcctct 120

ttccccgtgc agcagcctga gctctgcctc ctatgcaccc cagcaccccc atcagctccc 180

tcttctcctt caccagcccc gcagtgaagc ggctgctggg ctggaagcag ggagatgaag 240

aggagaagtg ggcagagaag gcggtggact ctttggtgaa gaagttaaag aagaagaaag 300

gcgccatgga tgaactggag agggcgctga gctgcccggg tcagcctagc aagtgtgtca 360

ccatcccacg gtccctcgat ggacgcctcc aggtgtccca ccgaaagggg ctgccccacg 420

tcatctactg ccgcgtgtgg cgctggccag acctgcagtc ccatcatgag ctgaagccct 480

tggagtgctg tgagttcccg ttcggctcca agcagaagga ggtctgcatc aacccatacc 540

attaccgcag agtggagacc ccagttctgc ctccagtgct ggtaccaaga cacagcgagt 600

acaaccctca gctcagcctc ctggccaagt tccgaagtgc ctcgctgcac agcgaacccc 660

tcatgccgca caacgccacc taccctgact ctttccagca gtctctctgt ccggcaccgc 720

cctcctcgcc aggccatgtg tttccgcagt ctccatgccc caccagctac ccgcactccc 780

ccggaagtcc ttccgagtca gacagtccct atcaacactc agacttccgg ccagtttgct 840

acgaggaacc ccagcactgg tgttctgttg cctactacga actaaacaac cgggtcggag 900

agactttcca ggcgtcctcg cggagcgtgc tcatagacgg cttcaccgac ccttccaata 960

acaggaatag gttttgcctt gggcttctct caaatgtaaa cagaaactcg accatagaaa 1020

acaccaggag gcacattgga aagggtgtgc atttgtacta cgttgggggc gaggtgtatg 1080

cggagtgcgt gagcgacagc agcatctttg tccagagccg gaactgcaac taccagcacg 1140

gcttccaccc ggccaccgtc tgcaagatcc ccagcggctg cagcctcaag gtcttcaaca 1200

accagctctt cgcccagctg ctcgcccagt ccgtgcacca cggctttgaa gtggtgtatg 1260

agctgacgaa gatgtgcacg attcggatga gctttgtgaa gggctgggga gcagagtatc 1320

atcgccagga tgtcacgagc acccccctgct ggatcgagat ccatcttcat ggaccgctgc 1380

agtggttgga taaggtgctc actcagatgg gctccccaca caaccctatc tcttcagtgt 1440

cttaagtcac gtcgtcagcc acgttgccac agaacagact cgggcagggg cttccatcgt 1500

ggcaaccgca gctaatgcag ggttccggat gcagatgtaa atacacgtgt aacgcatccg 1560

agtcacgttt atatcaccgt ttgttttgtg ctacctactt 1600

<210> 27
<211> 430
<212> PRT
<213> Mus musculus

<400> 27

Met His Pro Ser Thr Pro Ile Ser Ser Leu Phe Ser Phe Thr Ser Pro
1 5 10 15

Ala Val Lys Arg Leu Leu Gly Trp Lys Gln Gly Asp Glu Glu Glu Lys
20 25 30

Trp Ala Glu Lys Ala Val Asp Ser Leu Val Lys Lys Leu Lys Lys Lys
35 40 45

Lys Gly Ala Met Asp Glu Leu Glu Arg Ala Leu Ser Cys Pro Gly Gln
50 55 60

Pro Ser Lys Cys Val Thr Ile Pro Arg Ser Leu Asp Gly Arg Leu Gln
65 70 75 80

Val Ser His Arg Lys Gly Leu Pro His Val Ile Tyr Cys Arg Val Trp
85 90 95

Arg Trp Pro Asp Leu Gln Ser His His Glu Leu Lys Pro Leu Glu Cys
100 105 110

Cys Glu Phe Pro Phe Gly Ser Lys Gln Lys Glu Val Cys Ile Asn Pro
115 120 125

Tyr His Tyr Arg Arg Val Glu Thr Pro Val Leu Pro Pro Val Leu Val
130 135 140

Pro Arg His Ser Glu Tyr Asn Pro Gln Leu Ser Leu Leu Ala Lys Phe
145 150 155 160

```
Arg Ser Ala Ser Leu His Ser Glu Pro Leu Met Pro His Asn Ala Thr
                165                 170                 175

Tyr Pro Asp Ser Phe Gln Gln Ser Leu Cys Pro Ala Pro Pro Ser Ser
            180                 185                 190

Pro Gly His Val Phe Pro Gln Ser Pro Cys Pro Thr Ser Tyr Pro His
        195                 200                 205

Ser Pro Gly Ser Pro Ser Glu Ser Asp Ser Pro Tyr Gln His Ser Asp
    210                 215                 220

Phe Arg Pro Val Cys Tyr Glu Glu Pro Gln His Trp Cys Ser Val Ala
225                 230                 235                 240

Tyr Tyr Glu Leu Asn Asn Arg Val Gly Glu Thr Phe Gln Ala Ser Ser
                245                 250                 255

Arg Ser Val Leu Ile Asp Gly Phe Thr Asp Pro Ser Asn Asn Arg Asn
            260                 265                 270

Arg Phe Cys Leu Gly Leu Leu Ser Asn Val Asn Arg Asn Ser Thr Ile
        275                 280                 285

Glu Asn Thr Arg Arg His Ile Gly Lys Gly Val His Leu Tyr Tyr Val
    290                 295                 300

Gly Gly Glu Val Tyr Ala Glu Cys Val Ser Asp Ser Ser Ile Phe Val
305                 310                 315                 320

Gln Ser Arg Asn Cys Asn Tyr Gln His Gly Phe His Pro Ala Thr Val
                325                 330                 335

Cys Lys Ile Pro Ser Gly Cys Ser Leu Lys Val Phe Asn Asn Gln Leu
            340                 345                 350

Phe Ala Gln Leu Leu Ala Gln Ser Val His His Gly Phe Glu Val Val
        355                 360                 365

Tyr Glu Leu Thr Lys Met Cys Thr Ile Arg Met Ser Phe Val Lys Gly
    370                 375                 380

Trp Gly Ala Glu Tyr His Arg Gln Asp Val Thr Ser Thr Pro Cys Trp
385                 390                 395                 400
```

Ile Glu Ile His Leu His Gly Pro Leu Gln Trp Leu Asp Lys Val Leu
405 410 415

Thr Gln Met Gly Ser Pro His Asn Pro Ile Ser Ser Val Ser
420 425 430

INTERNATIONAL SEARCH REPORT

In I Application No
PCT/US2005/003229

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 A61K38/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	TAMAKI ET AL: "Intracellular signaling of osteogenic protein-1 through Smad5 activation" JOURNAL OF CELLULAR PHYSIOLOGY, vol. 177, 1998, pages 355-363, XP008048090 * See page 355 (Abstract) *	1-5,7-49
Y	AOKI ET AL: "Synergistic effects of different bone morphogenetic protein type I receptors on alkaline phosphatase induction" JOURNAL OF CELL SCIENCE, vol. 114, 2001, pages 1483-1489, XP002302367 * See pages 1483-1484 (Abstract and Introduction, esp. last sentence), figures 1-5, and pages 1487-1488 (Discussion) *	1-5,7-49

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 June 2005	17/06/2005

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Korsner, S-E

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	YEH ET AL: "Osteogenic protein-1 (OP-1, BMP-7) induces osteoblastic cell differentiation of the pluripotent mesenchymal cell line C2C12" JOURNAL OF CELLULAR BIOCHEMISTRY, vol. 87, 2002, pages 292-304, XP002302368 * See pages 300-303 (Discussion) *	1-5,7-49
Y	FORSLUND ET AL: "A comparative dose-response study of cartilage-derived morphogenetic protein (CDMP)-1, -2 and -3 for tendon healing in rats" JOURNAL OF ORTHOPAEDIC RESEARCH, vol. 21, 2003, pages 617-621, XP002330748 * See pages 617-618 (Introduction), and pages 619-621 (Discussion) *	1-5,7-49
Y	TAKASE ET AL: "Induction of Smad6 mRNA by bone morphogenetic proteins" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 244, 1998, pages 26-29, XP002219536 * See page 26 (Abstract) *	1-5,7-49
Y	MACÍAS-SILVA ET AL: "Specific activation of Smad1 signaling pathways by the BMP7 type I receptor, ALK2" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, 1998, pages 25628-25636, XP002330481 * See page 25630 (right column) - page 25631 (left column) *	1-5,7-49
Y	NISHIDA ET AL: "Osteogenic protein 1 stimulates cell-associated matrix assembly by normal human articular chondrocytes" ARTHRITIS & RHEUMATISM, vol. 43, 2000, pages 206-214, XP002330479 * See pages 206-207 (Introduction) *	1-5,7-49
Y	SIMANK ET AL: "Das Wachstumsfaktorkomposit aus GDF-5 und mineralisiertem Kollagen verbessert die Ausheilung einer Hüftkopfnekrose" ORTHOPÄDE, vol. 33, January 2004 (2004-01), pages 68-75, XP002330749 * See page 71 (English summary) *	1-5,7-49
Y	US 6 531 445 B1 (COHEN ET AL) 11 March 2003 (2003-03-11) * See columns 2-3 (tissues) and 6-7 (proteins useful) *	1-5,7-49

-/--

INTERNATIONAL SEARCH REPORT

Inte I Application No
PCT/US2005/003229

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 03/059932 A (YISSUM RESEARCH DEV. COMP. OF THE HEBREW UNIV. / GBF-GESELLSCHAFT) 24 July 2003 (2003-07-24) * See page 1 (Field of the invention), pages 11 and 21 *	1-5,7-49
P,X	YEH ET AL: "Cartilage-derived morphogenetic proteins enhance the osteogenic protein-1-induced osteoblastic cell differention of C2C12 cells" JOURNAL OF CELLULAR PHYSIOLOGY, vol. 201, December 2004 (2004-12), pages 401-408, XP008048134 * See page 401 (Abstract) *	1-5,7-49

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2005/003229

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although Claims 1-49 (partially) are directed to a method of treatment of the human body, the search has been carried out and based on the alleged effects of the compounds.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Inter Application No
PCT/US2005/003229

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6531445 B1	11-03-2003	US 2004102373 A1	27-05-2004
		US 6565843 B1	20-05-2003
		AT 192931 T	15-06-2000
		AU 669127 B2	30-05-1996
		AU 2564592 A	05-04-1993
		AU 670558 B2	25-07-1996
		AU 3176293 A	27-04-1993
		CA 2104678 A1	12-09-1992
		CA 2116559 A1	01-04-1993
		CA 2116562 A1	18-03-1993
		CA 2363965 A1	17-09-1992
		DE 69231062 D1	21-06-2000
		DE 69231062 T2	15-02-2001
		EP 0601106 A1	15-06-1994
		EP 0601135 A1	15-06-1994
		ES 2149776 T3	16-11-2000
		JP 6510989 T	08-12-1994
		JP 7502021 T	02-03-1995
		US 2003125230 A1	03-07-2003
		WO 9304692 A1	18-03-1993
		WO 9305751 A2	01-04-1993
		US 2003105004 A1	05-06-2003
		US 5656593 A	12-08-1997
		US 6333312 B1	25-12-2001
		US 6153583 A	28-11-2000
		US 2003104993 A1	05-06-2003
		US 6211146 B1	03-04-2001
		US 2003109445 A1	12-06-2003
		US 5650276 A	22-07-1997
		US 6022853 A	08-02-2000
		US 6395883 B1	28-05-2002
		US 5674844 A	07-10-1997
		US 6077823 A	20-06-2000
		US 5849686 A	15-12-1998
		US 5741641 A	21-04-1998
		US 6495513 B1	17-12-2002
		US 5834179 A	10-11-1998
		US 6399569 B1	04-06-2002
		US 5972884 A	26-10-1999
		US 5739107 A	14-04-1998
		US 6090776 A	18-07-2000
		US 5733878 A	31-03-1998
		US 2005065083 A1	24-03-2005
		US 2005096339 A1	05-05-2005
		US 6071708 A	06-06-2000
		US 5854071 A	29-12-1998
		US 5994131 A	30-11-1999
		US 2002049159 A1	25-04-2002
		US 6506729 B1	14-01-2003
WO 03059932 A	24-07-2003	AU 2002366974 A1	30-07-2003
		CA 2471891 A1	24-07-2003
		EP 1467624 A2	20-10-2004
		WO 03059932 A2	24-07-2003
		US 2003228292 A1	11-12-2003
		US 2003219423 A1	27-11-2003